

Array imaging of localized objects in homogeneous and heterogeneous media*

Anwei Chai^{†1}, Miguel Moscoso^{‡2}, and George Papanicolaou^{§3}

¹Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305 USA

²Gregorio Millán Institute, Universidad Carlos III de Madrid, Madrid 28911, Spain

³Department of Mathematics, Stanford University, Stanford, CA 94305 USA

Abstract

We present a comprehensive study of the resolution and stability properties of sparse promoting optimization theories applied to narrow band array imaging of localized scatterers. We consider homogeneous and heterogeneous media, and multiple and single scattering situations. When the media is homogeneous with strong multiple scattering between scatterers, we give a non-iterative formulation to find the locations and reflectivities of the scatterers from a nonlinear inverse problem in two steps, using either single or multiple illuminations. We further introduce an approach that uses the top singular vectors of the response matrix as optimal illuminations, which improves the robustness of sparse promoting optimization with respect to additive noise. When multiple scattering is negligible, the optimization problem becomes linear and can be reduced to a hybrid- ℓ_1 method when optimal illuminations are used. When the media is random, and the interaction with the unknown inhomogeneities can be primarily modeled by wavefront distortions, we address the statistical stability of these methods. We analyze the fluctuations of the images obtained with the hybrid- ℓ_1 method, and we show that it is stable with respect to different realizations of the random medium provided the imaging array is large enough. We compare the performance of the hybrid- ℓ_1 method in random media to the widely used Kirchhoff migration and the multiple signal classification methods.

AMS classification scheme numbers. 34B27, 78A46, 78A48

Keywords. array imaging, multiple scattering, random media, sparse promoting optimization, robustness, statistical stability

*This version: December 11, 2015.

[†]anwei@math.stanford.edu

[‡]moscoso@math.uc3m.es

[§]papanico@math.stanford.edu

1 Introduction

The recent mathematical theory of compressed sensing [18, 19, 20, 11, 12] has been shown to be very promising in a number of areas as diverse as medicine [29], biomedicine [37], geophysics [39], radar [3], astronomy [6], or microscopy [40]. Most inverse problems in these areas are considered to be underdetermined, meaning that we do not have unique solutions and, therefore, it is apparently impossible to identify which one is indeed the correct one. What makes compressed sensing at once interesting is that, often, the sought solution is known to be structured in the sense that it is sparse or compressible, which means that it depends upon a small number of parameters. This additional information changes the imaging problem dramatically because we can exploit the sparsity of the image and look for the simplest one that tends to be the right one.

In this paper, we study narrow-band, active array imaging of a small number of localized scatterers using both single and multiple illuminations. The goal is to determine the positions and reflectivities of the scatterers from the echoes recorded at an array of sensors when a few narrow band signals are sent to probe the medium. By localized scatterers we mean scatterers whose diameter is small compared to the wavelength. Hence, the Foldy-Lax approximation to the wave equation can be used to model wave propagation in the medium [26, 27, 28]. The number of scatterers is small because only a small portion of the region of interest is occupied by scatterers and, thus, the image we wish to recover is sparse. We study the case in which the interaction between the scatterers is strong so that multiple scattering is important, and the case in which the interaction is small so that multiple scattering is negligible. We consider imaging in homogeneous media and imaging in randomly inhomogeneous media with significant scattering from the inhomogeneities. We restrict this study to the case in which the full waveform at the array is available for imaging, which means that in the frequency domain both amplitudes and phases can be measured and recorded. For the case in which the phases cannot be recorded we refer to [15, 22, 23, 13, 14, 32, 31].

In this work we consider narrow-band array systems and, therefore, the frequency diversity content of the data measured at the array is very limited. There is an extensive literature on imaging techniques that deal with this problem. Kirchhoff migration [4], matched field imaging [1], and Multiple Signal Classification (MUSIC) [36] are among the most used techniques. As in narrow-band array imaging the data is scarce, and hence, there are infinitely many configurations of scatterers that match the data set, we formulate active array imaging as an optimization problem with constraints [16, 17]. When multiple scattering between the scatterers is important, the resulting problem is nonlinear, and therefore, it is apparently impossible to solve the optimization problem non-iteratively [17]. We show, however, that the nonlinearity can be avoided through a two-step process that effectively linearizes the inverse problem. In the first step, we treat the scatterers as equivalent sources and we recover their locations and strengths. In the second step, once the locations of the scatterers are fixed, we recover their true reflectivities using a known relationship between the source strengths and the scatterer reflectivities. This is an explicit relation that comes from the Foldy-Lax equations, given the scatterer locations and the illumination.

When multiple scattering is significant some scatterers may be obscured due to screening effects. Therefore, not all the scatterers may be recovered from data generated by a single illumination. Indeed, given an array illumination, multiple scattering may reduce the effec-

tive illumination at certain locations due to destructive interferences of secondary sources coming from all the scatterers. To mitigate this additional problem of multiple scattering, we will discuss the use of multiple illuminations. The resulting optimization problem with multiple illuminations will be formulated as a joint sparsity recovery problem where we seek an unknown matrix whose columns share the same support. Thus, we seek for solution vectors corresponding to different illuminations that have a common support but have possibly different nonzero values.

The key point of the proposed two-step approach is the possibility of exact recovery of the locations of the equivalent sources in the first step. We give conditions on the array imaging setup and the measurement noise level under which the locations of the sources can be recovered exactly. The uniqueness and stability of the solution are analyzed, showing that the errors are proportional to the amount of noise in the data with a proportionality factor that depends on the sparsity of the solution and the mutual coherence of the sensing matrix [17]. These conditions are given for general imaging configurations. Conditions on the resolution of the images that guarantee exact recovery in the paraxial regime are derived in [21] for scatterers whose range is known. A more general paraxial model with scatterers at different ranges from the array is considered in [10]. The interesting case of imaging scatterers with small off-grid displacements are studied in [24, 10]. In [24], a simple perturbation method is proposed to reduce the gridding error for off-grid scatterers. The authors in [10] also present a very nice discussion on how to interpret the results obtained with ℓ_1 minimization when modeling errors due to off-grid displacements are significant.

For the case in which multiple scattering can be ignored, we introduce a hybrid approach that combines the use of the singular value decomposition (SVD) of the data matrix with ℓ_1 minimization [16]. We use the top right singular vectors of this matrix as illumination vectors to collect the data. Then, we project the data onto the subspace spanned by the top left singular vectors to filter out the unnecessary data and the noise, and to reduce the dimension of the linear system. Finally, ℓ_1 optimization is applied to this reduced linear system to obtain the sparsest solution. This hybrid- ℓ_1 method turns out to be very useful when imaging in random media.

Imaging in random media is fundamentally different from imaging in homogeneous or smoothly varying media. When the medium is inhomogeneous we know, at best, the large scale, but we cannot know the small scale structure. Hence, when the small structure of a medium is important we model it as a random spatial process. In these cases, it is essential to take into consideration the statistical stability of the images, which refers to the robustness of the imaging methods with respect to different realizations of the medium. In fact, many of the usual imaging methods used in homogeneous (or smoothly varying) media fail, even for broadband signals and large arrays, because the images become noisy and change unpredictably with the detailed features of the fluctuations of the medium. This is the case, for example, of the images obtained with Kirchhoff migration that depend on the particular realization of the random medium, and thus, they become unstable. Statistical stability typically holds only in broadband, we refer to [7, 8, 9] for details. Here, we consider narrow-band array systems. We show that, in these cases, large arrays are essential to stabilize the images when the medium is random. In particular, we show that the hybrid- ℓ_1 method and MUSIC are efficient and robust when the arrays are large enough. We compare, using numerical simulations, the images obtained with these methods with those

obtained with Kirchhoff migration. The numerical simulations show that the hybrid- ℓ_1 method becomes stable faster than MUSIC as the array size increases. Kirchhoff migration is, as expected, unstable even for very large arrays.

The analysis of imaging in random media is done using a relatively simple random phase model for the effects of the random medium. This model characterizes wave propagation in the high-frequency regime in random media with weak fluctuations and small correlation lengths compared to the wavelength. It is widely used, for example, in adaptive optics to compensate for rapidly changing distortions in the received wavefronts due to the atmospheric turbulence caused by changing temperature and wind conditions.

The paper is organized as follows. In §2, we formulate the array imaging problem in homogeneous media when the multiple scattering is important. In §3, we describe the optimization methods that determine the locations and reflectivities of scatterers using a two-step non-iterative approach, with and without multiple illuminations. In §4, we consider the case where multiple scattering is negligible, and we give a hybrid- ℓ_1 method that improves the resolution of the image. In §5, we consider imaging in random media. Using the simple random phase model, we show that the hybrid- ℓ_1 and the MUSIC methods are statistically stable provided the arrays are large. The effectiveness of all the methods is illustrated in various numerical examples with comparisons to other imaging methods in each of the sections. Section 6 contains our conclusions. The proofs of the theoretical results are given in the appendix.

2 Data model

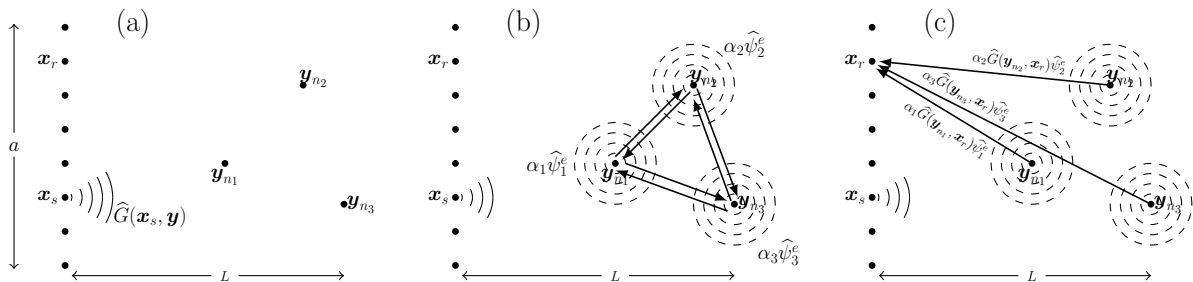


Figure 1: Schematic. (a) A linear array with $N = 8$ transducers probe a medium sending a spherical wave from x_s ; the illumination vector is $\hat{\mathbf{f}} = [0, 0, 0, 0, 0, \hat{f}_6, 0, 0]$. The medium is either homogeneous or inhomogeneous and the corresponding Green's function is $\hat{G}(\mathbf{x}, \mathbf{y})$. (b) There are $M = 3$ point-like scatterers with reflectivities α_j at positions y_{n_j} , $j = 1, 2, 3$. The interaction between them is strong creating effective (or secondary) sources $\gamma_{n_j} = \alpha_j \hat{\psi}_j^e$ that depend on all the scatterers' positions and their reflectivities. If those are known, the effective fields $\hat{\psi}_j^e$ can be computed by solving the Foldy-Lax equations described in Appendix A. (c) The response received at x_r is the superposition of all the scattered waves $\hat{\psi}_j^s = \alpha_j \hat{G}(y_{n_j}, x_r) \hat{\psi}_j^e$ from the scatterers.

Probing of the medium can be done with many different types of arrays, transmitters and recording devices. Also transducers that can both sense and transmit are usually employed.

Besides, the geometric layout of the arrays depend on the application (acoustics, seismology, radar, ...) and they may be arranged in a transmission or a backscattering configuration. To fix ideas, we will consider an active array consisting of N transducers located at positions \mathbf{x}_i , $i = 1, \dots, N$, placed in front of the medium to be probed (see Fig. 1 (a)). Here, and in the rest of the paper, we use boldface lower case letters for vectors, capital letters in boldface for matrices, and the corresponding letters, without boldface, for the entries of the matrices. To ensure that the transducers behave like an array of aperture $a = (N - 1)h$, and not like separate entities, they are separated by a distance h of the order of wavelength $\lambda = \frac{2\pi c_0}{\omega}$ of the probing signals, where c_0 is the wave speed in homogeneous medium and ω is the corresponding frequency.

We will assume that the object we wish to image consists of M randomly positioned point-like scatterers. The medium can be homogeneous or inhomogeneous. Multiple scattering among the scatterers may or may not be important. All the scatterers, with unknown reflectivities $\alpha_j \in \mathbb{C}$, where \mathbb{C} stands for the complex field, and positions \mathbf{y}_{n_j} , $j = 1, \dots, M$, are within a region of interest called the image window (IW), which is centered at a distance L from the array. We discretize the IW using a uniform grid of points \mathbf{y}_j , $j = 1, \dots, K$, and we introduce the true *reflectivity vector*

$$\boldsymbol{\rho}_0 = [\rho_{01}, \dots, \rho_{0K}]^T \in \mathbb{C}^K,$$

such that $\rho_{0k} = \sum_{j=1}^M \alpha_j \delta_{\mathbf{y}_{n_j}, \mathbf{y}_k}$, $k = 1, \dots, K$, where $\delta_{\cdot, \cdot}$ is the classical Kronecker delta and \cdot^T is the transpose only operation, while \cdot^* stands for the conjugate transpose. We further assume that each scatterer is located at one of the K grid points, so $\{\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}\} \subset \{\mathbf{y}_1, \dots, \mathbf{y}_K\}$. For a study of off-grid scatterers we refer to [24, 10].

To write the data received on the array in a compact form, we define the Green's function vector

$$\hat{\mathbf{g}}(\mathbf{y}; \omega) = [\hat{G}(\mathbf{x}_1, \mathbf{y}; \omega), \hat{G}(\mathbf{x}_2, \mathbf{y}; \omega), \dots, \hat{G}(\mathbf{x}_N, \mathbf{y}; \omega)]^T \quad (1)$$

at location \mathbf{y} in the IW, where $\hat{G}(\mathbf{x}, \mathbf{y}; \omega)$ denotes the free-space Green's function of the (homogeneous or inhomogeneous) medium that characterizes the propagation of a signal of angular frequency ω from point \mathbf{y} to point \mathbf{x} . When the medium is homogeneous,

$$\hat{G}(\mathbf{x}, \mathbf{y}; \omega) = \hat{G}_0(\mathbf{x}, \mathbf{y}; \omega) = \frac{\exp(i\kappa|\mathbf{x} - \mathbf{y}|)}{4\pi|\mathbf{x} - \mathbf{y}|}, \quad \kappa = \frac{\omega}{c_0}, \quad (2)$$

and we have the Green's function vector in a homogeneous medium as

$$\hat{\mathbf{g}}_0(\mathbf{y}; \omega) = [\hat{G}_0(\mathbf{x}_1, \mathbf{y}; \omega), \hat{G}_0(\mathbf{x}_2, \mathbf{y}; \omega), \dots, \hat{G}_0(\mathbf{x}_N, \mathbf{y}; \omega)]^T.$$

If $\hat{\mathbf{f}} = [\hat{f}_1, \dots, \hat{f}_N]^T$ is the illumination vector whose entries are the signals sent from the transmitters in the array, then $\hat{\mathbf{g}}(\mathbf{y}; \omega)\hat{\mathbf{f}}$ gives the field at position \mathbf{y} in a free-space.

We further introduce the $N \times K$ sensing matrix

$$\mathcal{G}(\omega) = [\hat{\mathbf{g}}(\mathbf{y}_1; \omega) \cdots \hat{\mathbf{g}}(\mathbf{y}_K; \omega)] \quad (3)$$

that maps a distribution of sources in the IW to the data received on the array. With this notation, the full response matrix, can be written as

$$\hat{\mathbf{P}} = \mathcal{G} \text{diag}(\boldsymbol{\rho}_0) \mathbf{Z}^{-1}(\boldsymbol{\rho}_0) \mathcal{G}^T = \mathcal{G} \text{diag}(\boldsymbol{\rho}_0) \mathcal{G}_{FL}^T(\boldsymbol{\rho}_0). \quad (4)$$

Here, and in all that follows, we drop the dependence of waves and measurements on the frequency ω . In (4), $\mathbf{Z}^{-1}(\boldsymbol{\rho}_0)$ denotes the inverse of the Foldy-Lax matrix which depends on the unknown reflectivity vector $\boldsymbol{\rho}_0$ (see Appendix A). To motivate (4), consider an illumination vector $\hat{\mathbf{f}} = [\hat{f}_1, \dots, \hat{f}_N]^T$ (see Fig. 1). Then, $\mathcal{G}_{FL}^T(\boldsymbol{\rho}_0)\hat{\mathbf{f}}$ gives the total field at each grid point of the IW, including multiple scattering between the scatterers and the interaction with the unknown inhomogeneities of the medium. The total field $\mathcal{G}_{FL}^T(\boldsymbol{\rho}_0)\hat{\mathbf{f}}$ is reflected by the scatterers on the grid that have reflectivities given by the vector $\boldsymbol{\rho}_0$, and then it is backpropagated to the array by the matrix \mathcal{G} . All the available information for imaging is contained in the array response matrix (4). If transmitters and receivers are located at the same positions, then (4) is symmetric.

For a fixed array configuration, wave propagation is completely described by the full response matrix $\hat{\mathbf{P}}$ (4). Indeed, the data received on the array due to an illumination vector $\hat{\mathbf{f}}$ is given by

$$\mathbf{b} = \hat{\mathbf{P}}\hat{\mathbf{f}}. \quad (5)$$

3 Active array imaging in homogeneous media

In this section, we formulate the inverse problem of active array imaging when the medium is homogeneous and, therefore, the wavefronts are not distorted. In this case, the response matrix can be written as

$$\hat{\mathbf{P}} = \mathcal{G}_0 \text{diag}(\boldsymbol{\rho}_0) \mathbf{Z}^{-1}(\boldsymbol{\rho}_0) \mathcal{G}_0^T = \mathcal{G}_0 \text{diag}(\boldsymbol{\rho}_0) \mathcal{G}_{0FL}^T(\boldsymbol{\rho}_0), \quad (6)$$

where $\mathcal{G}_0 = [\hat{\mathbf{g}}_0(\mathbf{y}_1) \cdots \hat{\mathbf{g}}_0(\mathbf{y}_K)]$ denotes the sensing matrix in a homogeneous medium, and $\mathbf{Z}^{-1}(\boldsymbol{\rho}_0)$ is the inverse of the Foldy-Lax matrix (53) with $\hat{G}(\mathbf{y}_i, \mathbf{y}_j) = \hat{G}_0(\mathbf{y}_i, \mathbf{y}_j)$ (see Appendix A). The object to be imaged is an ensemble of small but strong scatterers whose mutual interaction cannot be ignored. To determine their positions and reflectivities we use the collected data using a single illumination $\hat{\mathbf{f}}$ in subsection 3.1, and using multiple illuminations $\hat{\mathbf{f}}^j$, $j = 1, \dots, v$, in subsections 3.2 and 3.3. In signal processing literature, the corresponding problems are called *Single Measurement Vector (SMV)* and *Multiple Measurement Vector (MMV)* problems, respectively.

3.1 Imaging using single illumination

For a given illumination vector $\hat{\mathbf{f}}$, we define the operator $\mathcal{A}_{\hat{\mathbf{f}}}$ via

$$\mathcal{A}_{\hat{\mathbf{f}}}[\boldsymbol{\rho}_0]\boldsymbol{\rho}_0 = \hat{\mathbf{P}}\hat{\mathbf{f}},$$

which maps the reflectivity vector $\boldsymbol{\rho}_0$ to the data (5). From (6) it follows that

$$\mathcal{A}_{\hat{\mathbf{f}}}[\boldsymbol{\rho}] = [\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_1)\hat{\mathbf{g}}_0(\mathbf{y}_1) \cdots \hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_K)\hat{\mathbf{g}}_0(\mathbf{y}_K)], \quad (7)$$

where $\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_j) = \hat{\mathbf{g}}_{FL}^T(\mathbf{y}_j)\hat{\mathbf{f}}$, $j = 1, \dots, K$, are scalars meaning the total field at the grid points \mathbf{y}_j due to the illumination $\hat{\mathbf{f}}$, with $\hat{\mathbf{g}}_{FL}(\mathbf{y}_j)$ being the j^{th} column of the matrix \mathcal{G}_{0FL} .

Using this notation, active array imaging with a single illumination amounts to finding the unknown reflectivity vector $\boldsymbol{\rho}_0 \in \mathbb{C}^K$ from the system of N equations

$$\mathcal{A}_{\hat{\mathbf{f}}}[\boldsymbol{\rho}]\boldsymbol{\rho} = \mathbf{b}. \quad (8)$$

In a typical array imaging configuration, the number of transducers N is much smaller than the number of the grid points K in the IW and, hence, (8) is underdetermined. Furthermore, due to the multiple scattering, $\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_j)$, $j = 1, \dots, K$, depend on the unknown reflectivity vector $\boldsymbol{\rho}_0$, which makes (8) nonlinear with respect to $\boldsymbol{\rho}$. Such nonlinearity makes us think that non-iterative inversion is inapplicable to solve (8). However, by rearranging the terms in these equations, we can reformulate the problem to solve for the locations of scatterers directly, without any iteration, and then to recover the reflectivities of each scatterer in a second single step, as we explain next.

3.1.1 Support recovery

The localization problem is by far much more difficult than the estimation of reflectivities, which is a straightforward inversion if the former is exact. To localize the scatterers without any iteration, we introduce the *effective source vector*

$$\boldsymbol{\gamma}_{\hat{\mathbf{f}}} = \text{diag}(\boldsymbol{\rho})\mathbf{Z}^{-1}(\boldsymbol{\rho})\mathcal{G}_0^T \hat{\mathbf{f}}, \quad (9)$$

and seek for its support. Then, according to (6) $\mathcal{A}_{\hat{\mathbf{f}}}[\boldsymbol{\rho}]\boldsymbol{\rho} = \mathcal{G}_0\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$, and (8) becomes

$$\mathcal{G}_0\boldsymbol{\gamma}_{\hat{\mathbf{f}}} = \mathbf{b}, \quad (10)$$

which is linear for the new unknown $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$. Note that in the formulation (8) the operator $\mathcal{A}_{\hat{\mathbf{f}}}[\boldsymbol{\rho}]$ depends on the illumination $\hat{\mathbf{f}}$, whereas in (10) the unknown $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$ is the one which depends on $\hat{\mathbf{f}}$.

It is important to emphasize that due to the existence of multiple scattering, the solution of (10) may not give all the support of $\boldsymbol{\rho}_0$. This is not a flaw of the new formulation, but an implicit problem of array imaging when multiple scattering is important. Indeed, it is possible that one or several scalars $\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_j)$, $j = 1, \dots, K$, are very small or even zero, and hence, the corresponding scatterers become dark. This is the well-known *screening effect* which makes scatterers undetectable, and that is manifested in our formulation making some components of the effective source vector $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$ arbitrarily small.

Since (10) is underdetermined and the effective source vector $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$ is sparse because $M \ll N \ll K$, we solve the ℓ_1 minimization problem

$$\min \|\boldsymbol{\gamma}_{\hat{\mathbf{f}}}\|_{\ell_1} \quad \text{s.t.} \quad \mathcal{G}_0\boldsymbol{\gamma}_{\hat{\mathbf{f}}} = \mathbf{b} \quad (11)$$

when data is noiseless. When the data is contaminated by a noise vector \mathbf{e} with finite energy, we then solve the relaxed problem

$$\min \|\boldsymbol{\gamma}_{\hat{\mathbf{f}}}\|_{\ell_1} \quad \text{s.t.} \quad \|\mathcal{G}_0\boldsymbol{\gamma}_{\hat{\mathbf{f}}} - \mathbf{b}\|_{\ell_2} < \delta, \quad (12)$$

for some given positive constant δ . In (11) and (12), $\|\boldsymbol{\gamma}\|_{\ell_1} = \sum_{i=1}^K |\gamma_i|$ and $\|\boldsymbol{\gamma}\|_{\ell_2} = \sqrt{\sum_{i=1}^K |\gamma_i|^2}$.

The following theorem gives conditions under which (11) recovers the positions and the strengths of the effective sources exactly if the data is noiseless. The proof follows that in [16, 17].

Theorem 3.1. *Assume that the resolution of the IW is such that*

$$\max_{i \neq j} \left| \frac{\hat{\mathbf{g}}_0^*(\mathbf{y}_i) \hat{\mathbf{g}}_0(\mathbf{y}_j)}{\|\hat{\mathbf{g}}_0(\mathbf{y}_i)\|_{\ell_2} \|\hat{\mathbf{g}}_0(\mathbf{y}_j)\|_{\ell_2}} \right| < \epsilon. \quad (13)$$

If the number of effective sources M is such that $\epsilon M < 1/2$, then $\boldsymbol{\gamma}_{0\hat{\mathbf{f}}} = \text{diag}(\boldsymbol{\rho}_0) \mathbf{Z}^{-1}(\boldsymbol{\rho}_0) \mathcal{G}_0^T \hat{\mathbf{f}}$ is the unique solution to (11) with support fully contained by that of $\boldsymbol{\rho}_0$.

Remark 3.2. *The condition (13) of the sensing matrix \mathcal{G}_0 is determined by the array imaging configuration. It is a measure of how linearly independent the columns of the sensing matrix are and it is related to the array geometry as shown in [17].*

Remark 3.3. *It turns out that to prove the result in Theorem 3.1, the condition (13) has to be satisfied only on the support of $\boldsymbol{\gamma}_{0\hat{\mathbf{f}}}$. This means that if the distance between the effective sources is known a priori to be large so (13) holds for the set of indices corresponding to its support, the discretization of the IW can be as small as we want.*

The next theorem provides an important stability result for the problem (12), the proof of which is given by Theorem 4.3 in [17].

Theorem 3.4. *For a given array imaging configuration, with resolution condition given by (13), the solution $\boldsymbol{\gamma}_{*\hat{\mathbf{f}}}$ to (12) satisfies*

$$\|\boldsymbol{\gamma}_{*\hat{\mathbf{f}}} - \boldsymbol{\gamma}_{0\hat{\mathbf{f}}}\|_{\ell_2} \leq \frac{\delta}{\sqrt{1 - (M-1)\epsilon}}, \quad (14)$$

provided $\delta \geq \|\mathbf{e}\|_{\ell_2} \sqrt{1 + \frac{M(1-(M-1)\epsilon)}{(1-2M\epsilon+\epsilon)^2}}$, where \mathbf{e} is the noise vector added to the data. Moreover, the support of $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$ is fully contained in that of $\boldsymbol{\gamma}_{0\hat{\mathbf{f}}}$, and all the components satisfying*

$$|(\boldsymbol{\gamma}_{0\hat{\mathbf{f}}})_j| > \delta / \sqrt{1 - (M-1)\epsilon} \quad (15)$$

are within the support of $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$.*

3.1.2 Reflectivity estimation

Optimization (11) (or (12)) gives the effective source vector $\boldsymbol{\gamma}_{\hat{\mathbf{f}}}$. In a second step, we compute the true reflectivities from the solution of this problem at once. Let Λ_* be the support of the recovered solution such that $|\Lambda_*| = M' \leq M$, and $\boldsymbol{\gamma}_{\hat{\mathbf{f}}, M'}$ be the solution vector on Λ_* . From (9), we have

$$\boldsymbol{\gamma}_{\hat{\mathbf{f}}, M'} = \text{diag}(\mathbf{Z}^{-1}(\boldsymbol{\rho}_{M'}) \mathcal{G}_0^T \hat{\mathbf{f}}) \boldsymbol{\rho}_{M'} = \text{diag}(\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_{n_1}), \dots, \hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_{n_{M'}})) \boldsymbol{\rho}_{M'},$$

where $\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_{n_j}) = \hat{\mathbf{g}}_{0FL}^T(\mathbf{y}_{n_j}) \hat{\mathbf{f}}$. Note that the scalars $\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_{n_j})$ are the exciting fields at the scatterers' positions, and that the effective sources γ_{n_j} are the true reflectivities ρ_{n_j} multiplied

by the exciting fields (see Fig. 1 (b)). Hence, using (50) we can compute $\widehat{g}_{\widehat{\mathbf{f}}}(\mathbf{y}_{n_j})$ explicitly as follows

$$\widehat{g}_{\widehat{\mathbf{f}}}(\mathbf{y}_{n_j}) = \widehat{\mathbf{g}}_0^T(\mathbf{y}_{n_j})\widehat{\mathbf{f}} + \sum_{k=1, k \neq j}^{M'} \gamma_k \widehat{G}_0(\mathbf{y}_{n_j}, \mathbf{y}_{n_k}), \quad j = 1, \dots, M'. \quad (16)$$

Then, the true reflectivities of the scatterers are recovered by

$$\rho_{n_j} = \gamma_{n_j} / \widehat{g}_{\widehat{\mathbf{f}}}(\mathbf{y}_{n_j}), \quad j = 1, \dots, M'. \quad (17)$$

When the data contains additive noise, we choose the support Λ_\star of the solution recovered by (12) such that all the components of $\boldsymbol{\gamma}_{\widehat{\mathbf{f}}, M'}$ satisfy (15).

3.2 Imaging using multiple arbitrary illuminations

Imaging with a single illumination can be very sensitive to additive noise, especially when the noise level is high. Moreover, the *screening effect* can cause the failure of recovering some of the scatterers. Note that, for a fixed imaging configuration, the *screening effect* depends on the illumination vector \mathbf{f} and the amount of noise in the data. When the effective source at \mathbf{y}_j is below the noise level because $\widehat{g}_{\widehat{\mathbf{f}}}(\mathbf{y}_j)$ is small, the corresponding scatterer cannot be detected. This motivates us to consider active array imaging with multiple illuminations. In this case, active array imaging is modeled as a joint sparsity recovery problem, in which we seek for a matrix solution \mathbf{X} whose columns share the same support. By increasing the diversity of illuminations, we are able to minimize the *screening effect* and have higher chance of locating all the scatterers more stably.

3.2.1 Support recovery

Instead of simply stacking multiple data vectors \mathbf{b}^j obtained from different illuminations $\widehat{\mathbf{f}}^j$, $j = 1, \dots, v$, and solving the corresponding augmented linear system as in §3.1, we formulate the problem using the *MMV* approach where the unknown vectors $\boldsymbol{\gamma}^j$ corresponding to each illumination $\widehat{\mathbf{f}}^j$ are arranged into a matrix. More precisely, let $\mathbf{B} = [\mathbf{b}^1, \dots, \mathbf{b}^v]$ be the matrix whose columns are the data vectors generated by all the illuminations, and $\mathbf{X} = [\boldsymbol{\gamma}^1, \dots, \boldsymbol{\gamma}^v]$ be the unknown matrix whose j^{th} column corresponds to the *effective source vector* $\boldsymbol{\gamma}^j$ under illumination $\widehat{\mathbf{f}}^j$. Thus, we formulate the problem of active array imaging with multiple measurements as solving the matrix-matrix equation

$$\mathcal{G}_0 \mathbf{X} = \mathbf{B} \quad (18)$$

for \mathbf{X} . The sparsity of the matrix variable \mathbf{X} is characterized by the number of nonzero rows. Thus, we define the row-support of a given matrix \mathbf{X} by

$$\text{rowsupp}(\mathbf{X}) = \{i : \exists j \text{ s.t. } X_{ij} \neq 0\}.$$

When the matrix \mathbf{X} degenerates to a column vector, the row-support reduces to the support of that vector. Similar to the ℓ_1 norm relaxation used in the SMV problem in §3.1, the sparsest solution using multiple illuminations is given by the solution to the convex problem

$$\min \Xi_1(\mathbf{X}) \quad \text{s.t.} \quad \mathcal{G}_0 \mathbf{X} = \mathbf{B}, \quad (19)$$

where Ξ_1 is a convenient convex relaxation of the size of $\text{row}\text{supp}(\mathbf{X})$. As in [17], we take

$$\Xi_1(\mathbf{X}) = \sum_{i=1}^N \|X_{i\cdot}\|_{\ell_2}, \quad (20)$$

where $X_{i\cdot}$ is the i^{th} row of the matrix. When the data vectors are contaminated with additive noise vectors \mathbf{e}^j , $j = 1, \dots, v$, we have the matrix-matrix equation

$$\mathcal{G}_0 \mathbf{X} = \mathbf{B} + \mathcal{E}, \quad (21)$$

where $\mathcal{E} = [\mathbf{e}^1 \dots \mathbf{e}^v]$, and we seek a solution to

$$\min \Xi_1(\mathbf{X}) \quad \text{s.t.} \quad \|\mathcal{G}_0 \mathbf{X} - \mathbf{B}\|_F < \delta, \quad (22)$$

for a pre-specified constant δ , where the Frobenius norm is given by

$$\|\mathbf{X}\|_F = \left(\sum_{i=1}^N \sum_{j=1}^K |X_{ij}|^2 \right)^{\frac{1}{2}}.$$

Remark 3.5. *Similar to Theorems 3.1 and 3.4 for array imaging with a single illumination, we have results regarding the exact recovery and stability of the solution to (19) and (22) for imaging using multiple illuminations. These results are proved in [17]. We note, however, that these results do not provide a quantitative improvement when the number of measurements v increases. Intuitively, this lack of improvement is justifiable since the measurements obtained from randomly chosen illuminations could be rather ineffective. There is no guarantee that randomly picked illuminations bring more information useful for imaging. In practice, however, we do observe a general improvement in the images formed with multiple random illuminations. In §3.3, we use selective illuminations, obtained from the SVD of the array response matrix $\hat{\mathbf{P}}$ to increase the efficiency of the MMV approach.*

3.2.2 Reflectivity estimation

Once we obtain the matrix \mathbf{X}_\star from (19) (or (22)), whose columns are the effective sources corresponding to the different illuminations, we compute in a second step the true reflectivities as follows. For each component i in the support such that the stability condition in Theorem 4.3 in [17] is satisfied, we compute the reflectivities ρ_i^j corresponding to each illumination j by applying (16) and (17). Then, we take the average $\frac{1}{v} \sum_{j=1}^v \rho_i^j$ as the estimated reflectivity.

3.3 Imaging using optimal illuminations

To further improve the efficiency, and especially the robustness of (22), we propose to use selective illuminations as multiple illumination vectors [17]. The optimal set of illumination vectors can be obtained systematically from the SVD of the full response matrix $\hat{\mathbf{P}}$. If

the full array response matrix is not available, they can be obtained from an iterative time reversal process as discussed in [16]. Let the SVD of $\hat{\mathbf{P}}$ be

$$\hat{\mathbf{P}} = \hat{\mathbf{U}} \mathbf{\Sigma} \hat{\mathbf{V}}^* = \sum_{j=1}^{\tilde{M}} \sigma_j \hat{\mathbf{U}}_{.j} \hat{\mathbf{V}}_{.j}^*,$$

where $\hat{\mathbf{U}}_{.j}$ and $\hat{\mathbf{V}}_{.j}$ are the left and right singular vectors, respectively, and the nonzero singular values σ_j are given in descending order as $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\tilde{M}} > 0$, with $\tilde{M} \geq M$. When there is no noise in the data, $\tilde{M} = M$. Now, let the illumination vectors be the top right singular vectors, i.e. $\hat{\mathbf{f}}^j = \hat{\mathbf{V}}_{.j}$, $j = 1, \dots, v \leq \tilde{M}$. Then, we have

$$\mathbf{G}_0 \mathbf{X}_V = \mathbf{B}_{opt} = \hat{\mathbf{P}} \hat{\mathbf{V}}_{:,1:v} = [\sigma_1 \hat{\mathbf{U}}_{.1} \dots \sigma_v \hat{\mathbf{U}}_{.v}] + \tilde{\mathbf{E}}, \quad (23)$$

where the data matrix \mathbf{B}_{opt} contains all the essential information for imaging the scatterers. The improvement of the efficiency, when using optimal illuminations, comes from the fact that illuminations using singular vectors deliver most of the energy around the scatterers, even when multiple scattering is non-negligible. Therefore, taking a few top singular vectors is enough to focus around the scatterers that contribute to the data received on the imaging array.

3.4 Numerical experiments

We now present numerical simulations in two dimensions. In all the simulations shown below with a single illumination, we use the iterative shrinkage-thresholding algorithm GelMa proposed in [30] due to its flexibility with respect to the choice of the regularization parameter used in the algorithm. In the simulations with multiple illuminations we use an extension of this algorithm described in [17].

Figure 2 shows five scatterers placed within an IW of size $41\lambda \times 41\lambda$, which is at a distance $L = 100\lambda$ from a linear array with 100 transducers that are one wavelength λ apart (the spatial units of all the images is λ). The amplitudes of the reflectivities of the scatterers $|\alpha_j|$ are 2.96, 2.76, 2.05, 1.54 and 1.35. Their phases are set randomly in each realization. Note that for a given illumination $\hat{\mathbf{f}}$ and a scatterer configuration $\boldsymbol{\rho}_0$ with fixed amplitudes, the amount of multiple scattering depends on the realization of the phases in $\boldsymbol{\rho}_0$. For the amplitudes of the reflectivities chosen here, the amount of multiple scattering over single scattering ranges between 50% and 100%.

Figure 3 shows the images obtained by ℓ_1 norm minimization with 0% (left), 10% (middle), and 20% noise (right) when a single illumination probes the medium. When there is no noise in the data, ℓ_1 norm minimization recovers the positions and reflectivities of the scatterers exactly (the true locations are indicated with small white dots). However, when the data is corrupted by 10% and 20% of additive noise (middle and right images of Figure 3), ℓ_1 norm minimization fails. Some ghosts appear in the images and some scatterers are missing because of *screening effect*.

When multiple illuminations are used we expect a general improvement in the images. Figure 4 shows the results of the MMV algorithm when 5 random illuminations are used. By multiple random illuminations we mean a set of illuminations coming, each one, from

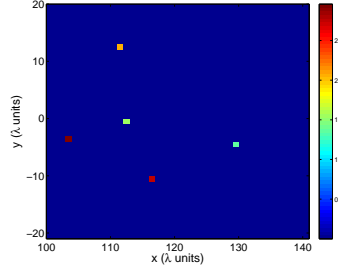


Figure 2: Original configuration of the scatterers in a 41×41 image window with grid points separated by 1.

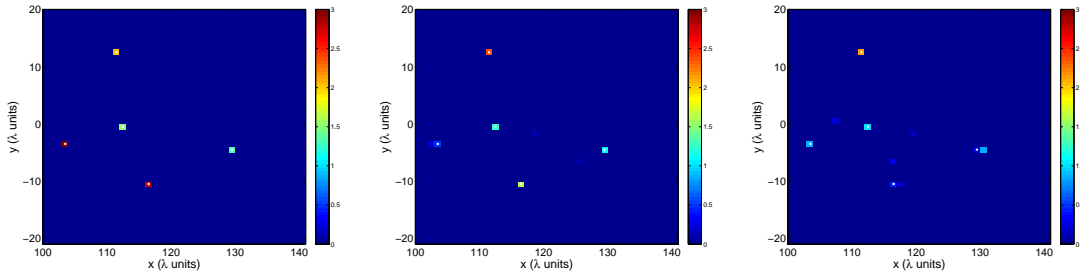


Figure 3: Images reconstructed by solving (11) and (12) when single illumination is used. From left to right, there is 0%, 10%, and 20% additive noise in the data.

only one of the transducers on the array at a time, i.e., $\hat{f}_p = 1$ and $\hat{f}_q = 0$ for $q \neq p$, with p chosen randomly at a time. The set of illuminations used in each image of Figure 4 is different. The data contain 10% (left), 20% (middle) and 50% (right) of additive noise. We observe that the images obtained with multiple illuminations are more stable with respect to additive noise. However, we still see some missing scatterers in the right image (50% of noise), and therefore, for this particular set of randomly chosen illuminations, imaging is still affected by *screening effect*.

Figure 4 shows the necessity of using “good” illuminations that maximize the information content of the data, especially when the signal-to-noise ratio (SNR) is low. Figure 5 displays the images obtained with 1 (left), 2 (middle), and 3 (right), optimal illuminations associated

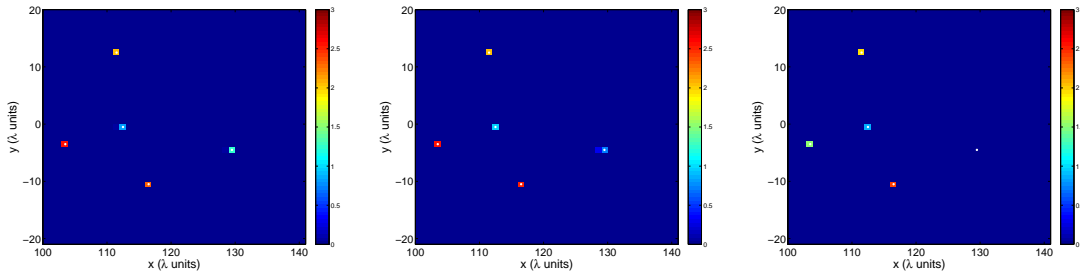


Figure 4: Images reconstructed by solving (22) when 5 random illuminations are used. From left to right, there is 10%, 20%, and 50% noise in the data.

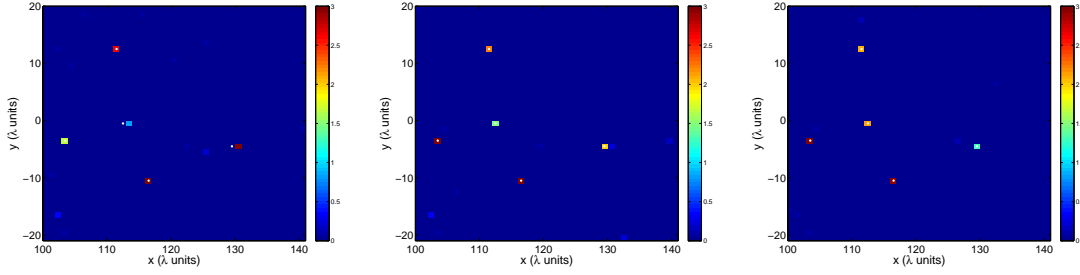


Figure 5: Images reconstructed by solving (22) when optimal illuminations are used. From left to right, the images are reconstructed by using 1, 2, and 3 top singular vectors as illumination vectors. There is 50% noise in the data.

to the singular vectors $\hat{\mathbf{f}}^j = \hat{\mathbf{V}}_{\cdot j}$, with $j = 1, 2, 3$. It is remarkable that only a few of them (2 or 3) are enough to obtain very good images, even with 50% of additive noise in the data.

4 Single scattering as a special case: The hybrid- ℓ_1 method

In situations where the scatterers are weak and/or are very far apart, the incident wave undergoes only one scattering event before coming back to the array. In these cases, the inverse of the Foldy-Lax matrix $\mathbf{Z}^{-1}(\boldsymbol{\rho})$ in (6) becomes the identity matrix, and the array response matrix under the so called Born approximation becomes

$$\hat{\mathbf{\Pi}} = \mathcal{G}_0 \text{diag}(\boldsymbol{\rho}_0) \mathcal{G}_0^T.$$

Then, the data received on the array when an illumination vector $\hat{\mathbf{f}}$ probes the medium is given by $\mathbf{b} = \hat{\mathbf{\Pi}} \hat{\mathbf{f}}$. Similar to (8), we can model the data \mathbf{b} as applying the operator $\mathcal{A}_{\hat{\mathbf{f}}}$ to the unknown reflectivity vector $\boldsymbol{\rho}_0$. Because multiple scattering is negligible, $\hat{g}_{\hat{\mathbf{f}}}(\mathbf{y}_j) = \hat{g}_0^T(\mathbf{y}_j) \hat{\mathbf{f}}$ in (7), and thus, (8) is now linear in $\boldsymbol{\rho}_0$, which we denote as $\mathcal{A}_{0\hat{\mathbf{f}}}$. Therefore, array imaging of localized scatterers in homogeneous media when multiple scattering is negligible amounts to solving the linear optimization problem

$$\min \|\boldsymbol{\rho}\|_{\ell_1} \quad \text{s.t.} \quad \mathcal{A}_{0\hat{\mathbf{f}}} \boldsymbol{\rho} = \mathbf{b}. \quad (24)$$

In (24) we assume that the data is noiseless. If the data contains noise, we solve

$$\min \|\boldsymbol{\rho}\|_{\ell_1} \quad \text{s.t.} \quad \|\mathcal{A}_{0\hat{\mathbf{f}}} \boldsymbol{\rho} - \mathbf{b}\|_{\ell_2} < \delta. \quad (25)$$

Because of the linearity of the equations, a single-step process is sufficient to recover both the locations and reflectivities of the scatterers from (24) or (25). Moreover, when multiple scattering is negligible, there is no *screening effect*, and hence, solving the set of equations generated from a single illumination alone can produce reliable images. Note that Theorems 3.1 and 3.4 can be applied to problems (24) and (25) as well.

Similarly to the case where multiple scattering was important, we can use multiple illuminations and the MMV approach to find the reflectivity vector $\boldsymbol{\rho}_0$. Furthermore, imaging

using optimal illuminations has a simpler form and weaker requirements on the imaging configuration when multiple scattering is negligible. Indeed, assume that the data is noiseless, and let the SVD of array response matrix be $\hat{\mathbf{\Pi}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^*$, with M singular values $\{\sigma_j\}_{j=1}^M$ greater than 0. Following the MMV formulation in §3.2, we have $\mathbf{b}^j = \hat{\mathbf{\Pi}}\hat{\mathbf{V}}_{\cdot j} = \sigma_j\hat{\mathbf{U}}_{\cdot j}$ for $j = 1, \dots, M$. On the other hand, $\mathbf{b}^j = \mathcal{A}_{0\hat{\mathbf{V}}_{\cdot j}}\boldsymbol{\rho}_0 = \mathcal{G}_0\boldsymbol{\gamma}_0^j$, where $\boldsymbol{\gamma}_0^j = \text{diag}(\hat{\mathbf{g}}_{\hat{\mathbf{V}}_{\cdot j}}(\mathbf{y}_k))\boldsymbol{\rho}_0$ is linear in $\boldsymbol{\rho}_0$, which is the major difference from (9). Now consider (23) with $\tilde{\mathcal{E}} = \mathbf{0}$ for simplicity, and project \mathbf{B}_{opt} onto the space spanned by the top left singular vectors corresponding to the significant singular values. Denoting $\mathbf{U} = [\hat{\mathbf{U}}_{\cdot 1}, \dots, \hat{\mathbf{U}}_{\cdot M}]$ and $\mathbf{X}_V = [\boldsymbol{\gamma}^1, \dots, \boldsymbol{\gamma}^M]$, we get

$$\begin{aligned} \mathbf{U}^*\mathcal{G}_0\mathbf{X}_V &= \begin{bmatrix} \hat{\mathbf{U}}_{\cdot 1}^*\mathcal{G}_0\boldsymbol{\gamma}^1 & \hat{\mathbf{U}}_{\cdot 1}^*\mathcal{G}_0\boldsymbol{\gamma}^2 & \cdots & \hat{\mathbf{U}}_{\cdot 1}^*\mathcal{G}_0\boldsymbol{\gamma}^M \\ \hat{\mathbf{U}}_{\cdot 2}^*\mathcal{G}_0\boldsymbol{\gamma}^1 & \hat{\mathbf{U}}_{\cdot 2}^*\mathcal{G}_0\boldsymbol{\gamma}^2 & \cdots & \hat{\mathbf{U}}_{\cdot 2}^*\mathcal{G}_0\boldsymbol{\gamma}^M \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{U}}_{\cdot M}^*\mathcal{G}_0\boldsymbol{\gamma}^1 & \hat{\mathbf{U}}_{\cdot M}^*\mathcal{G}_0\boldsymbol{\gamma}^2 & \cdots & \hat{\mathbf{U}}_{\cdot M}^*\mathcal{G}_0\boldsymbol{\gamma}^M \end{bmatrix} \\ &= \mathbf{U}^*\mathbf{B}_{opt} = \text{diag}(\sigma_1, \dots, \sigma_M). \end{aligned} \quad (26)$$

In (26), there are M unknowns $\boldsymbol{\gamma}^1, \dots, \boldsymbol{\gamma}^M$, and M^2 equations. However, when multiple scattering is negligible we can associate a nonzero singular value to each scatterer and the corresponding singular vector is given by

$$\hat{\mathbf{U}}_{\cdot j} = \overline{\hat{\mathbf{V}}_{\cdot j}} \propto \frac{\hat{\mathbf{g}}_0(\mathbf{y}_j)}{\|\hat{\mathbf{g}}_0(\mathbf{y}_j)\|_{\ell_2}}, \quad j = 1, \dots, M, \quad (27)$$

up to an arbitrary complex constant of modulus 1. Therefore, the rank-1 matrices $\hat{\mathbf{V}}_{\cdot j}\hat{\mathbf{U}}_{\cdot i}^*$ have a (non-trivial) eigenvalue different from zero only when $i = j$. That is, only the diagonal blocks of the matrix $\mathbf{U}^*\mathcal{G}_0\mathbf{X}_V$ contribute in (26), and hence, we can simplify (26) further as a set of M equations. Using (27) and

$$\hat{\mathbf{U}}_{\cdot i}^*\mathcal{G}_0\boldsymbol{\gamma}^j = \sum_{k=1}^N (\hat{\mathbf{U}}_{\cdot i}^*\hat{\mathbf{g}}_0(\mathbf{y}_k))(\hat{\mathbf{g}}_0(\mathbf{y}_k)^*\hat{\mathbf{V}}_{\cdot j})\rho_{0k} = \sum_{k=1}^N \hat{\mathbf{g}}_0^*(\mathbf{y}_k)(\hat{\mathbf{V}}_{\cdot j}\hat{\mathbf{U}}_{\cdot i}^*)\hat{\mathbf{g}}_0(\mathbf{y}_k)\rho_{0k},$$

we can write (26) as a set of M equations linear in $\boldsymbol{\rho}_0$ such that

$$\mathcal{B}\boldsymbol{\rho}_0 = [\sigma_1, \dots, \sigma_M]^T, \quad (28)$$

where

$$\mathcal{B} = \begin{bmatrix} \overline{\hat{\mathbf{g}}_0^*(\mathbf{y}_1)}\hat{\mathbf{U}}_{\cdot 1}\hat{\mathbf{g}}_0^T(\mathbf{y}_1)\hat{\mathbf{V}}_{\cdot 1} & \cdots & \overline{\hat{\mathbf{g}}_0^*(\mathbf{y}_K)}\hat{\mathbf{U}}_{\cdot 1}\hat{\mathbf{g}}_0^T(\mathbf{y}_K)\hat{\mathbf{V}}_{\cdot 1} \\ \cdots & \ddots & \cdots \\ \overline{\hat{\mathbf{g}}_0^*(\mathbf{y}_1)}\hat{\mathbf{U}}_{\cdot M}\hat{\mathbf{g}}_0^T(\mathbf{y}_1)\hat{\mathbf{V}}_{\cdot M} & \cdots & \overline{\hat{\mathbf{g}}_0^*(\mathbf{y}_K)}\hat{\mathbf{U}}_{\cdot M}\hat{\mathbf{g}}_0^T(\mathbf{y}_K)\hat{\mathbf{V}}_{\cdot M} \end{bmatrix}.$$

Therefore, the hybrid- ℓ_1 method is to solve for $\boldsymbol{\rho}_0$ from the SMV problem

$$\min \|\boldsymbol{\rho}\|_{\ell_1} \quad \text{s.t.} \quad \mathcal{B}\boldsymbol{\rho} = [\sigma_1, \dots, \sigma_M]^T. \quad (29)$$

The advantage of hybrid- ℓ_1 method compared to the MMV formulation is its simpler form and lower dimensionality. Moreover, the hybrid ℓ_1 method guarantees exact recovery, the sufficient condition for which also implies that it can form images with higher resolution (see [16] for details).

Theorem 4.1. *Assume that the scatterers are far part such that (27) is satisfied. Let \mathbf{E} be the submatrix of \mathbf{B} formed by normalized columns corresponding to the scatterers' location and \mathbf{S} be the submatrix formed by the remaining normalized columns of \mathbf{B} . If $\|\mathbf{S}\|_{1 \rightarrow 1} < 1 - \|\mathbf{E} - \mathbf{I}\|_{1 \rightarrow 1}$, where $\|\cdot\|_{1 \rightarrow 1}$ is the matrix 1-norm induced by the vector ℓ_1 norm, and \mathbf{I} is the identity matrix, then $\boldsymbol{\rho}_0$ is the unique solution to (29).*

4.1 Numerical experiments

Now we present the results of a numerical experiment that shows the performance of the hybrid- ℓ_1 method compared to those obtained with KM and MUSIC. We consider the same imaging set-up as in subsection 3.4. KM is a simple and robust imaging method (with respect to additive noise) which propagates the imaging data received on the array back to the medium. Under the setup of imaging in §2, the KM imaging functional can be written as

$$\boldsymbol{\rho}_{KM} = \mathcal{A}_{0\hat{f}}^* \mathbf{b}. \quad (30)$$

MUSIC, on the other hand, is a subspace projection algorithm that uses the SVD of the array response matrix. The idea of MUSIC is to project the reference illumination vectors $\hat{\mathbf{g}}_0(\mathbf{y}_j)$, $j = 1, \dots, M$, onto the noise space using the projection operator

$$\mathcal{P}\hat{\mathbf{g}}_0(\mathbf{y}) = \sum_{j=1}^M (\hat{U}_{\cdot j}^* \hat{\mathbf{g}}_0(\mathbf{y})) \hat{U}_{\cdot j} - \hat{\mathbf{g}}_0(\mathbf{y}). \quad (31)$$

Then, according to (27), the normalized functional

$$\mathcal{I}_{MUSIC}(\mathbf{y}_k) = \frac{\min_{1 \leq j \leq K} \|\mathcal{P}\hat{\mathbf{g}}_0(\mathbf{y}_j)\|_{\ell_2}}{\|\mathcal{P}\hat{\mathbf{g}}_0(\mathbf{y}_k)\|_{\ell_2}}, \quad k = 1, \dots, K, \quad (32)$$

will peak at the search points \mathbf{y}_k only at locations where there is a scatterer. Imaging with MUSIC (32) only locates of the scatterer's positions. Their reflectivities are usually obtained using a separate procedure after the locations are known.

Figure 6 shows the simulation results with 100% of additive noise in the data. From left to right, we display the images obtained with KM, MUSIC and the hybrid- ℓ_1 method. In the simulations for MUSIC and hybrid- ℓ_1 we assume that the top 5 singular vectors have been obtained. In the simulations for KM we use 5 random illuminations. It is apparent that the image created by the hybrid- ℓ_1 method is better than the other two. The hybrid- ℓ_1 method forms a clear image with perfect recovery of the location of the scatterers, and accurate estimates of their reflectivities. The robustness of the hybrid- ℓ_1 method relies on the selection of the top singular values so that the noise in the subspace formed by singular vectors associated with small singular values is filtered out.

5 Active array imaging in random media

Many natural media vary randomly in space, and therefore, waves propagating in these media also fluctuate in space. Thus the data collected on the array inherit the uncertainty of the

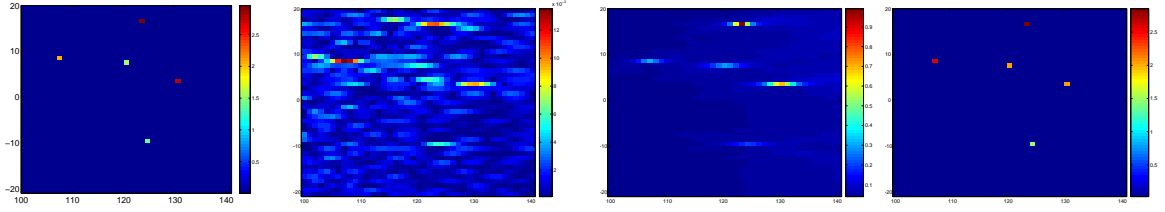


Figure 6: Images obtained using KM (left), MUSIC (middle), and hybrid- ℓ_1 (right) with 100% noise in data. We use 5 illuminations in KM.

fluctuations of the medium resulting in wave distortion. Because wave distortion effects induced by the inhomogeneities are very different from additive uncorrelated noise, imaging through random media is very different from that in homogeneous media. For example, the images obtained by Kirchhoff migration become noisy and change unpredictably with the detailed features of the medium, and thus, are useless. In this section, we use a simple random phase model to analyze the performance of the hybrid- ℓ_1 and MUSIC methods in random media. Both methods use the SVD of the response matrix. The random phase model distorts the wavefronts but keeps the amplitude of the waves unchanged and is valid in the regime of geometrical optics.

5.1 Random phase model

In random media, the Green's function that characterizes the wave propagation from \mathbf{x} to \mathbf{y} is given by the wave equation

$$\Delta \hat{G}(\mathbf{x}, \mathbf{y}) + \kappa^2 n^2(\mathbf{x}) \hat{G}(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), \quad (33)$$

where $n(\mathbf{x}) = \frac{c_0}{c(\mathbf{x})}$ is the random index of refraction of the medium with local wave speed $c(\mathbf{x})$. In a homogeneous medium, we have $c(\mathbf{x}) \equiv c_0$ for any location \mathbf{x} , and hence, $\hat{G}_0(\mathbf{x}, \mathbf{y})$ defined in (2) is the solution to (33). In random media, however, the wave speed $c(\mathbf{x})$ depends on the position \mathbf{x} . In this paper, we consider random fluctuations of the wave speed satisfying the model

$$\frac{1}{c^2(\mathbf{x})} = \frac{1}{c_0^2} \left(1 + \sigma \mu\left(\frac{\mathbf{x}}{l}\right) \right), \quad (34)$$

where l is the correlation length of the medium, σ determines the strength of the fluctuation around the constant speed c_0 , and $\mu(\cdot)$ is a stationary random process with zero mean and normalized autocorrelation function $R(|\mathbf{x} - \mathbf{x}'|) = \mathbb{E}(\mu(\mathbf{x})\mu(\mathbf{x}'))$, so $R(0) = 1$. Here, we only consider weak fluctuations such that $\sigma \ll 1$.

When the index of refraction $n(\mathbf{x})$ is random, there is no analytic solution to (33), and hence, numerical schemes are often used to obtain an approximated numerical solution. This can be time consuming, especially for large distances L . Instead, the random phase model provides an analytical approximation for the Green's function between two points at a distance of order $L \gg l \gg \lambda$ from each other, given by

$$\hat{G}(\mathbf{x}, \mathbf{y}) = \hat{G}_0(\mathbf{x}, \mathbf{y}) \exp \left[i\sigma\kappa |\mathbf{x} - \mathbf{y}| \int_0^1 \mu\left(\frac{\mathbf{x}}{l} + \frac{s}{l}(\mathbf{y} - \mathbf{x})\right) ds \right]. \quad (35)$$

This approximation is valid when (i) the wavelength λ is much smaller than the correlation length l so the geometric optics approximation holds, (ii) the correlation length l is much smaller than the propagation distance L so the statistics of the phase is Gaussian, and (iii) the strength of the fluctuation σ is small so the amplitude of the wave is kept unchanged, but large enough to ensure that the perturbations of the phases are not negligible. The last condition holds when $\frac{\sigma^2 L^3}{l^3} \ll \frac{\lambda^2}{\sigma^2 l L} \leq 1$ (for details, see [38, 35, 9]). Note that although we take weak fluctuations, the distortion of the wavefronts by the inhomogeneities of the medium is observable because the wave travels long distances.

Comparing (35) to the homogeneous Green's function (2) we see that, in this regime, only the phase is perturbed by the random medium and the magnitude remains unchanged. The following result shows that the second order moment of (35) is close to the expected value.

Proposition 5.1. *Assume that the autocorrelation function $R \in \mathbb{L}_2(\mathbb{R}_+)$ is differentiable and its derivative \dot{R} satisfies $\frac{\dot{R}(t)}{t} \in \mathbb{L}(\mathbb{R}_+)$ with exponential decay. Let \mathbf{y}_1 and \mathbf{y}_2 be two points on the same plane at a distance L from point \mathbf{x} , such that $\lambda \ll |\mathbf{y}_1 - \mathbf{y}_2| \ll L$. Then, for the Green's function (35) we have*

$$\mathbb{E} \left(\widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)} \right) \approx \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2}, \quad (36)$$

with

$$a_e = \sigma L \left(-1 - \frac{2L}{3l} \int_0^\infty \frac{\dot{R}(t)}{t} dt \right)^{\frac{1}{2}}. \quad (37)$$

According to Proposition 5.1, the back-propagated signal in random media is equal to that in a homogeneous medium times a Gaussian factor with variance $\frac{L^2}{\kappa^2 a_e^2}$. The length a_e is an intrinsic property of the random media. It depends on the propagation distance L and the statistics of the random fluctuations of the medium. According to (36), as a_e increases, the refocused spot size for time reversal in random media is tighter. Using the moment estimate (36), we immediately have the following variance estimate.

Corollary 5.2. *Under the same conditions given in Proposition 5.1, we have for (35)*

$$\mathbb{E} \left| \widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)} - \mathbb{E} \left(\widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)} \right) \right|^2 \approx \frac{1 - e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2}}{16\pi^2 |\mathbf{x} - \mathbf{y}_1|^2 |\mathbf{x} - \mathbf{y}_2|^2}. \quad (38)$$

This property of the Green's function (35) is generally referred to as *self-averaging*, which means that a time-reversed, back-propagated signal in a random medium refocuses near the source independently of the particular realization of the random medium. Self-averaging properties of back-propagated signals in random media have been shown under broad-band settings, see for example [25, 5, 2, 33, 34], but are not true, in general, for narrow-band signals. Corollary 5.2 states that the Green's function corresponding to the phase random model (35) exhibits the same properties, even for narrow-band pulses.

The next result states that a single realization of a back-propagated signal via the Green's function vector $\widehat{\mathbf{g}}(\mathbf{y})$ (1) is stable when the aperture of the imaging array becomes infinity, in the sense that its value is close to the back-propagated signal averaged over multiple realizations.

Proposition 5.3. *For a large aperture size a , the signal sent from \mathbf{y}_1 , recorded at the array, and back-propagated at \mathbf{y}_2 is statistically stable for (35) in the sense that*

$$\frac{\mathbb{E}|\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2))|^2}{\mathbb{E}\|\widehat{\mathbf{g}}(\mathbf{y}_1)\|^2\mathbb{E}\|\widehat{\mathbf{g}}(\mathbf{y}_2)\|^2} \rightarrow 0, \quad a \rightarrow \infty. \quad (39)$$

Furthermore, as shown in Appendix B, the decay rate of (39) is bounded by

$$\left(1 - e^{-\frac{\kappa^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2 a_e^2}\right) \left(\frac{l^2}{L^2 \log\left(1 + \frac{a^2}{4L^2}\right)}\right) \quad (40)$$

for any aperture size a . Thus, for short distances so L is of order a , the decay rate is only logarithmic in a . However, when the distance is large such that $L \gg a$ (e.g. the remote sensing regime), we can use a linear approximation on the logarithm function and obtain, up to a constant, the approximated decay rate

$$\left(1 - e^{-\frac{\kappa^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2 a_e^2}\right) \left(\frac{l}{a}\right)^2, \quad (41)$$

which implies a quadratic decay rate in a . Such quadratic decay rate in a can also be justified by using the paraxial approximation when $L \gg a$. As it is noted in Remark B.2, the decay rate in the paraxial approximation is given by (60), so the bound in (41) coincides with the paraxial approximation up to oscillations caused by the sinc function.

We emphasize that, although (40) indicates little improvement in the refocused spot size when the physical aperture of the array a is large, Proposition 5.3 implies that large arrays stabilize refocusing of narrow-band pulses in random media. The fact that large arrays improve the quality of imaging in cluttered media has been recognized for long time in radar and seismic imaging. The theoretical results obtained here using the random phase model support this empirical observation.

Next, we give a result that is essential for the performance of hybrid- ℓ_1 and MUSIC methods in random media. When imaging in random media, the Green's function vector $\widehat{\mathbf{g}}(\mathbf{y})$ is random and not known. The best we can do is to use the deterministic Green's function vector $\widehat{\mathbf{g}}_0(\mathbf{y})$ which is, in general, quite different from the random one. Furthermore, replacement of $\widehat{\mathbf{g}}(\mathbf{y})$ by $\widehat{\mathbf{g}}_0(\mathbf{y})$ might give good results for some realizations of the random medium but not for others. The next result tells us that a single realization of a narrow-band backpropagated signal $\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2)$, with the Green's function corresponding to the phase random model, is stable as the aperture of the array becomes infinity, which means that its value is close to the back-propagated signal averaged over multiple realizations.

Proposition 5.4. *Let \mathbf{y}_1 and \mathbf{y}_2 be two points satisfying $\lambda \ll |\mathbf{y}_1 - \mathbf{y}_2| \ll L$. Then, for the Green's function (35), we asymptotically have*

$$\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) \rightarrow 0, \quad a \rightarrow \infty, \quad (42)$$

and for the mixed inner products

$$\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) \rightarrow 0, \quad a \rightarrow \infty, \quad (43)$$

where the limit is under probability measure \mathbb{P} induced by $\mu(\cdot)$. Moreover, the mixed inner product is also statistically stable in the sense that

$$\frac{\mathbb{E} |\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2))|^2}{\|\widehat{\mathbf{g}}_0(\mathbf{y}_1)\|^2 \mathbb{E}\|\widehat{\mathbf{g}}(\mathbf{y}_2)\|^2} \rightarrow 0, \quad a \rightarrow \infty. \quad (44)$$

Proposition 5.4 implies that imaging in random media, i.e., a narrow-band pulse $\widehat{\mathbf{g}}(\mathbf{y}_2)$ back-propagated via $\widehat{\mathbf{g}}_0(\mathbf{y}_1)$ in the medium, is statistically stable for large arrays, and the pulse heading to the point \mathbf{y}_1 will focus around the correct location \mathbf{y}_2 .

5.2 The hybrid- ℓ_1 method

Assume that multiple scattering between the scatterers is negligible, and wave distortion is well described by the random phase model (35). Then, the response matrix has the form

$$\widehat{\mathbf{P}} = \sum_{j=1}^M \alpha_j \widehat{\mathbf{g}}(\mathbf{y}_{n_j}) \widehat{\mathbf{g}}^T(\mathbf{y}_{n_j}), \quad (45)$$

with random Green's function vectors $\widehat{\mathbf{g}}(\mathbf{y}) = [\widehat{G}(\mathbf{x}_1, \mathbf{y}), \dots, \widehat{G}(\mathbf{x}_N, \mathbf{y})]^T$, where $\widehat{G}(\mathbf{x}, \mathbf{y})$ is given by (35). Because multiple scattering is not important, the transformation $\mathbf{A}_{\widehat{\mathbf{f}}} \boldsymbol{\rho} = \mathbf{b}$ that relates the reflectivity vector $\boldsymbol{\rho}_0$ with the data measured at the array \mathbf{b} is linear. However, when the medium contains random inhomogeneities, this linear transformation is random because the column vectors of the operator $\mathbf{A}_{\widehat{\mathbf{f}}}$ are random functions. Therefore, to solve for the reflectivity vector $\boldsymbol{\rho}_0$, we have to approximate the unknown Green's function vectors $\widehat{\mathbf{g}}(\mathbf{y})$ by the ones for a homogeneous medium, after which, $\mathbf{A}_{\widehat{\mathbf{f}}}$ becomes the linear operator $\mathbf{A}_{0\widehat{\mathbf{f}}}$ given in §4. The substitution of $\widehat{\mathbf{g}}(\mathbf{y})$ by $\widehat{\mathbf{g}}_0(\mathbf{y})$ introduces a discrepancy between model and data, and thus, we must solve the ℓ_1 minimization problem (25) with an inequality constraint when data from a single illumination is available. We note that to estimate the error bound of such discrepancy is nontrivial due to the existing correlation in the measurement noise in random media.

According to Proposition 5.4, when the scatterers are far apart, the random Green's function vectors $\widehat{\mathbf{g}}(\mathbf{y}_{n_j})$, $j = 1, \dots, M$, are approximately orthogonal in probability as the size of the imaging array becomes large. In this case, we can associate to each of the M scatterers a nonzero singular value $\sigma_j = \alpha_j \|\widehat{\mathbf{g}}(\mathbf{y}_{n_j})\|_{\ell_2}^2$ with singular vectors

$$\widehat{U}_{\cdot j} = \frac{\widehat{\mathbf{g}}(\mathbf{y}_{n_j})}{\|\widehat{\mathbf{g}}(\mathbf{y}_{n_j})\|_{\ell_2}}, \quad \widehat{V}_{\cdot j} = \frac{\overline{\widehat{\mathbf{g}}(\mathbf{y}_{n_j})}}{\|\widehat{\mathbf{g}}(\mathbf{y}_{n_j})\|_{\ell_2}}, \quad j = 1, \dots, M. \quad (46)$$

If the full array response matrix $\widehat{\mathbf{P}}$ (or its SVD) is available, we can use the hybrid- ℓ_1 as follows. We use the right singular vectors of $\widehat{\mathbf{P}}$ as the illumination vectors, and we project the data to the space spanned by the left singular vectors in the same way as shown in §4. Thus, we form the same hybrid- ℓ_1 optimization given in (29) and there is no need to estimate the error bound of the discrepancy between model and data. Recall that $(ij)^{\text{th}}$ entry of the hybrid- ℓ_1 matrix \mathbf{B} is related to the singular vectors in the form of $\widehat{\mathbf{g}}_0^*(\mathbf{y}_j) \widehat{U}_{\cdot i} \widehat{\mathbf{g}}_0^T(\mathbf{y}_j) \widehat{V}_{\cdot i}$, and the singular vectors of the response matrix $\widehat{\mathbf{P}}$ are related to the random Green's function

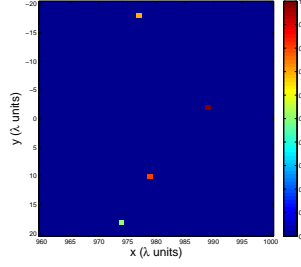


Figure 7: Original configuration of the scatterers in a 41×41 image window with grid points separated by 1.

vectors as in (46). Hence, each entry is the mixed inner product (43) between $\hat{\mathbf{g}}_0(\mathbf{y}_j)$ and $\hat{\mathbf{g}}(\mathbf{y}_i)$.

Without loss of generality, assume that the M scatterers correspond to the first M columns in \mathbf{B} , i.e. $n_j = j$, $j = 1, \dots, M$. We can write \mathbf{B} in the form of block matrices as $\mathbf{B} = [\mathbf{B}_{M \times M} \mathbf{S}]$, where the columns in \mathbf{S} correspond to grid points where there is no scatterers. Due to Proposition 5.4, the submatrix $\mathbf{B}_{M \times M}$ is a diagonal matrix \mathbf{D} perturbed by a matrix \mathbf{E} with $\|\mathbf{E}\|_{1 \rightarrow 1} \ll 1$. The value of $\|\mathbf{E}\|_{1 \rightarrow 1}$ depends on the minimal distance between any two scatterers, and is close to zero if they are well separated. Due to the incomplete phase cancellation in the mixed inner products, \mathbf{D} is composed of complex values but, as long as the scatterers are far apart, $\mathbf{D} + \mathbf{E}$ is diagonal dominated. Furthermore, because of the statistical stability of the mixed inner product (44), if $\|\mathbf{D}^{-1}\mathbf{S}\|_{1 \rightarrow 1} < 1 - \|\mathbf{D}^{-1}\mathbf{E}\|_{1 \rightarrow 1}$, (29) gives the right solution even though there exists (correlated) noise in the data caused by the random media. This saves computational time by eliminating the need for difficult error bounds testing.

5.3 MUSIC

In random media, the exact knowledge of the reference illumination $\hat{\mathbf{g}}(\mathbf{y})$ is unknown, and therefore imaging using MUSIC amounts to approximate the unknown illumination with that in homogeneous context, i.e.

$$\mathcal{I}_{MUSIC}^{RM}(\mathbf{y}_k) = \frac{\min_{1 \leq j \leq K} \|\mathcal{P}\hat{\mathbf{g}}_0(\mathbf{y}_j)\|_{\ell_2}}{\|\mathcal{P}\hat{\mathbf{g}}_0(\mathbf{y}_k)\|_{\ell_2}}, \quad k = 1, \dots, K. \quad (47)$$

According to (46) and (43), the proxy $\hat{\mathbf{g}}_0(\mathbf{y})$ will focus around the scatterers when the imaging array is large enough. However, the image obtained by (47) will have lower resolution compared to the ideal case of (32), since the random phase in $\hat{\mathbf{g}}(\mathbf{y})$ cannot be fully cancelled out using the homogenous Green's function vector $\hat{\mathbf{g}}_0(\mathbf{y})$.

5.4 Numerical experiments

We now present numerical simulations in two dimensions to illustrate the performance of the MUSIC and hybrid- ℓ_1 imaging methods in random media. We consider a random medium with correlation length $l = 20\lambda$, and standard deviation of the fluctuations $\sigma = 0.1\%$. For

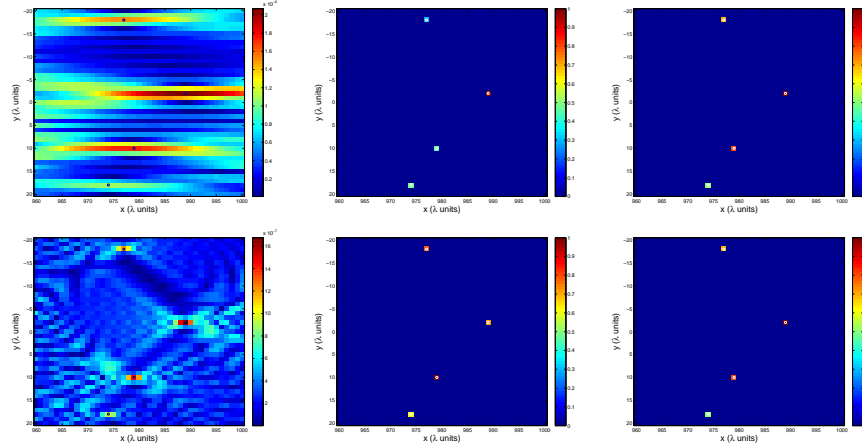


Figure 8: Top row: small array. Bottom row: large array. KM (left column), MUSIC (center column) and hybrid- ℓ_1 (right column) images in homogeneous media. There is no additive noise in the data.

comparison purposes we also show images obtained by Kirchhoff migration with a single illumination sent from the central transducer of the array.

We consider a small and a large array, both consisting of 501 transducers uniformly distributed over the aperture. The small array has an aperture of $25l$, and the large array an aperture of $100l$. Four scatterers are placed within an IW of size $41\lambda \times 41\lambda$, which is at a distance $L = 50l$ from the linear array, see Figure 7. We discretize the IW using a uniform grid with points separated by one wavelength λ (the spatial unit in all the figures is λ). The amplitudes of the reflectivities of the scatterers, $|\alpha_j|$, are 0.8, 1.0, 0.5, and 0.7. The phases are set randomly in each realization. We compute the array data (45) using the Green's function given by (35). The line integral of the random field in (35) is approximated by a quadrature rule.

As a reference, we show in Figure 8 the images obtained in a homogeneous medium using KM (left column), MUSIC (center column), and hybrid- ℓ_1 methods (right column) with noiseless data. The top and bottom rows show the results for the small array and the large array, respectively. As expected, the resolution of the KM images improves greatly for large arrays. On the other hand, in a homogeneous medium, MUSIC and hybrid- ℓ_1 achieve an excellent resolution, even for small arrays. The images shown in Figure 8 do not change too much when the data is corrupted with up to 100% of additive noise [16].

The situation changes when there is correlated noise in the data because the signals propagate through a random medium with a complex structure. This is illustrated in Figure 9, where we show the images produced by these imaging methods using a small array ($a = 25l$). The three imaging methods show different behaviors though. Kirchhoff migration completely fails to image the scatterers, as can be seen in the top row the figure. There is not only degradation in the resolution, but also loss of stability. Observe that the images obtained with KM are significantly different from one realization to another.

The images obtained with MUSIC (middle row of Figure 9) are also blurred compared to those obtained in a homogeneous medium. Furthermore, the images change from one

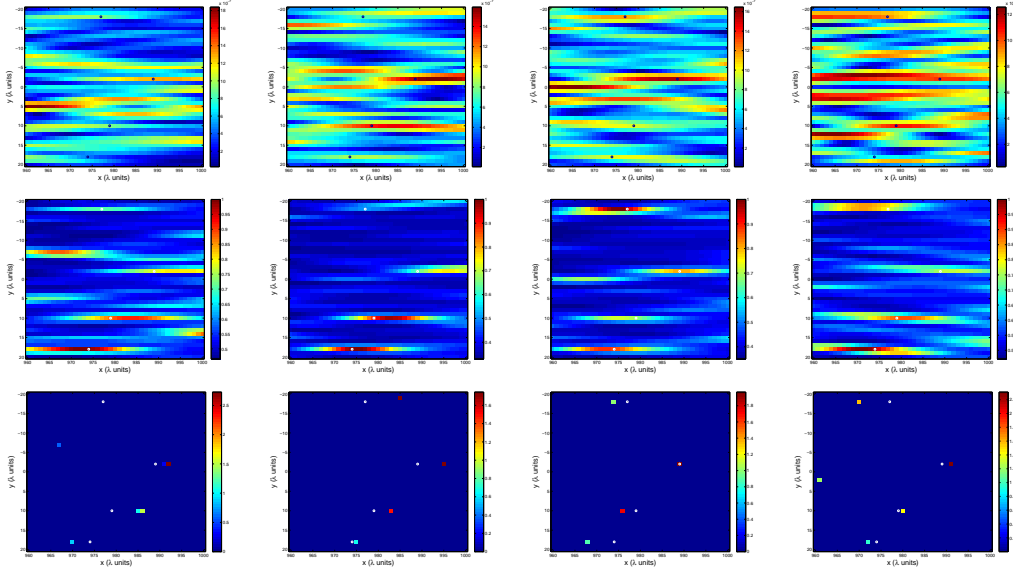


Figure 9: **Small array.** KM (top row) , MUSIC (center row) and hybrid- ℓ_1 (bottom row) images in the same four realizations of a random medium.

realization of the random medium to another and, therefore, MUSIC is also unstable if the array size is small. The hybrid- ℓ_1 method also produces images that change from one realization to another (bottom row of Figure 9), but tries to keep a good resolution to provide a sparse solution. Observe that the detected scatterers dance along the cross-range direction around the true locations indicated in the figure with white dots. The images obtained with the hybrid- ℓ_1 method also show some ghosts.

These problems are overcome when the array is large ($a = 100l$), see Figure 10. As expected, the resolution of the images obtained by KM and MUSIC improve a lot. However, KM still fails to image in random media as it produces clutter noise in the images from which it is hard or impossible to identify the location of the four scatterers. On the other hand, MUSIC and the hybrid- ℓ_1 method are able to recover the sparse solution. We observe, though, that the performance of the hybrid- ℓ_1 method is better. The locations of the four scatterers are found exactly by this method.

6 Conclusions

We present a comprehensive study of optimization based methods applied to narrow band array imaging of localized scatterers. We have considered homogeneous and heterogeneous media. When the media is homogeneous but multiple scattering between the scatterers is important, we give a non-iterative formulation of the nonlinear inverse problem that allows us to determine the locations and reflectivities of the scatterers non-iteratively using sparsity promoting optimization. We also propose to apply optimal illuminations to improve the robustness of the imaging methods and the resolution of the images. When multiple scattering is negligible, the optimization problem becomes linear. In this case, our formulation can

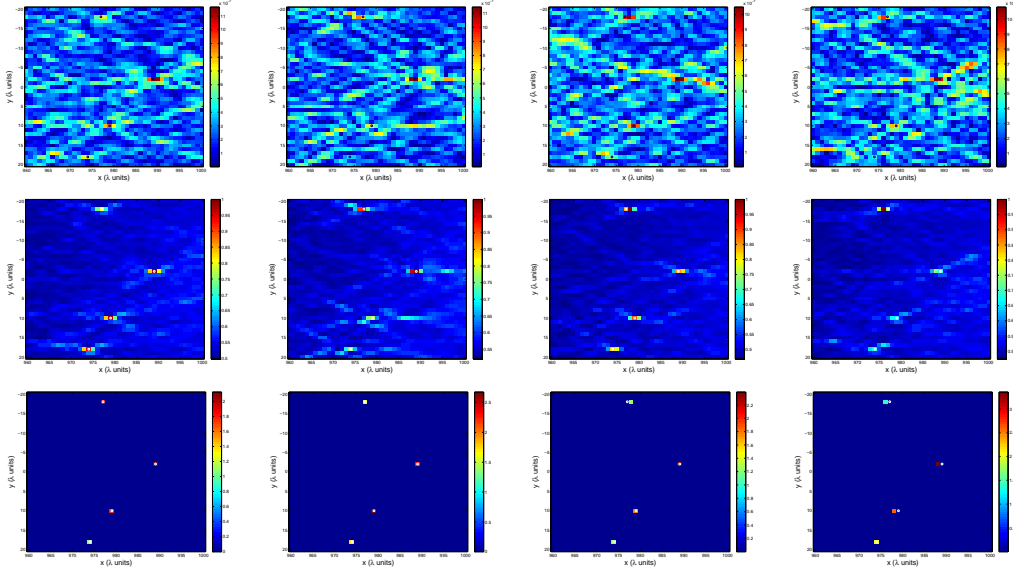


Figure 10: **Large array.** KM (top row) , MUSIC (center row) and hybrid- ℓ_1 (bottom row) images in the same four realizations of a random medium.

be reduced to a hybrid- ℓ_1 method that uses the optimal illuminations and ℓ_1 minimization. This method reduces the dimensionality of the problem, filters out the noise, and keeps all the essential properties of ℓ_1 minimization.

When the media is random, we study the important concept of statistical stability which relates to the robustness of the imaging methods with respect to different realizations of the random media. Provided the imaging array is large enough, we show that the hybrid- ℓ_1 method gives very accurate results and is statistically stable.

We illustrate the theoretical results with various numerical examples and compared the performance of the proposed optimization based methods to the widely used Kirchhoff migration and the MUSIC methods.

Acknowledgments

AC's work was partially supported by a Hewlett Packard Stanford Graduate Fellowship. MM's work was partially supported by the Spanish MICINN grant FIS2013-41802-R. GP's work was partially supported by AFOSR grant FA9550-14-1-0275.

Appendix

A Foldy-Lax model

Multiple scattering between point-like scatterers is modeled by means of the Foldy-Lax equations [26, 27, 28]. Under this framework, the scattered wave received at transducer \mathbf{x}_r

due to a narrow band signal of angular frequency ω sent from \mathbf{x}_s can be written as the superposition of all scattered waves from the M scatterers at $\{\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}\}$ (see Fig. 1 (c)), so

$$\hat{P}(\mathbf{x}_r, \mathbf{x}_s) = \sum_{j=1}^M \hat{\psi}_j^s(\mathbf{x}_r; \mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}). \quad (48)$$

Here, $\hat{\psi}_j^s(\mathbf{x}_r; \mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M})$ represents the scattered wave observed at \mathbf{x}_r due to the emanating wave from the scatterer at position \mathbf{y}_{n_j} . It depends on the positions of all the scatterers and is given by

$$\hat{\psi}_j^s(\mathbf{x}_r; \mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}) = \alpha_j \hat{G}(\mathbf{x}_r, \mathbf{y}_{n_j}) \hat{\psi}_j^e(\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}), \quad (49)$$

where $\hat{\psi}_j^e(\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M})$ represents the exciting field at the scatterer located at \mathbf{y}_{n_j} . Ignoring the self-interacting fields, the exciting field at \mathbf{y}_{n_j} is equal to the sum of the incident field $\hat{\psi}_j^{inc} := \hat{\psi}^{inc}(\mathbf{y}_{n_j}, \mathbf{x}_s)$ at \mathbf{y}_{n_j} and the scattered fields at \mathbf{y}_{n_j} due to all scatterers except for the one at \mathbf{y}_{n_j} . Hence, it is given by

$$\hat{\psi}_j^e(\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}) = \hat{\psi}^{inc}(\mathbf{y}_{n_j}, \mathbf{x}_s) + \sum_{k \neq j} \alpha_k \hat{G}(\mathbf{y}_{n_j}, \mathbf{y}_{n_k}) \hat{\psi}_k^e(\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}), \quad (50)$$

for $j = 1, 2, \dots, M$. This is a self-consistent system of M equations for the M unknown exciting fields

$$\hat{\psi}_1^e := \hat{\psi}_1^e(\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}), \dots, \hat{\psi}_M^e := \hat{\psi}_M^e(\mathbf{y}_{n_1}, \dots, \mathbf{y}_{n_M}).$$

We write (50) in matrix form as

$$\mathbf{Z}_M(\boldsymbol{\alpha}) \boldsymbol{\Phi}^e = \boldsymbol{\Phi}^{inc}, \quad (51)$$

where $\boldsymbol{\Phi}^e = [\hat{\psi}_1^e, \dots, \hat{\psi}_M^e]^T$ and $\boldsymbol{\Phi}^{inc} = [\hat{\psi}_1^{inc}, \dots, \hat{\psi}_M^{inc}]^T$ are vectors whose components are the exciting and incident fields on the M scatterers, respectively, and

$$(Z_M(\boldsymbol{\alpha}))_{ij} = \begin{cases} 1 & i = j \\ -\alpha_j \hat{G}(\mathbf{y}_{n_i}, \mathbf{y}_{n_j}) & i \neq j \end{cases} \quad (52)$$

is the $M \times M$ Foldy-Lax matrix which depends on the reflectivities $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_M]^T$, and on the positions of the scatterers \mathbf{y}_{n_j} . With the solution of (51), we use (49) and (48) to compute the scattered data received at the array.

We extend the $M \times M$ matrix $\mathbf{Z}_M(\boldsymbol{\alpha})$, defined only for pairwise combinations of scatterers at \mathbf{y}_{n_j} , to a larger $K \times K$ matrix

$$(Z(\boldsymbol{\rho}_0))_{ij} = \begin{cases} 1, & i = j \\ -\rho_{0j} \hat{G}(\mathbf{y}_i, \mathbf{y}_j), & i \neq j \end{cases} \quad (53)$$

which includes all pairwise combinations of the K grid points \mathbf{y}_j in the IW. Using (53) and (3), we can write the response matrix as (4).

B Proof of results in §5

In order to prove Proposition 5.1, we need to show the following lemma first.

Lemma B.1. *Assume the autocorrelation function R of the random field in $\mathbb{L}_2(\mathbb{R}_+)$ with derivative \dot{R} such that $\frac{\dot{R}(t)}{t} \in \mathbb{L}(\mathbb{R}_+)$ and both R and \dot{R} decay exponentially. Then*

$$\int_0^\infty \frac{\dot{R}(t)}{t} dt < 0.$$

Proof. Since $R \in \mathbb{L}_2(\mathbb{R}_+)$, the Fourier transform exists, i.e.

$$\widehat{R}(|\mathbf{k}|) = \int_{\mathbb{R}^3} R(|\mathbf{x}|) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}.$$

Moreover, $\widehat{R}(\cdot) \geq 0$. Using spherical coordinate (r, ϕ, θ) we obtain

$$\begin{aligned} \widehat{R}(|\mathbf{k}|) &= \int_0^{2\pi} \int_0^\pi \int_0^\infty R(r) r^2 \sin \phi e^{i|\mathbf{k}|r \cos \phi} dr d\phi d\theta \\ &= 2\pi \int_0^\infty R(r) r^2 \int_0^\pi \sin \phi e^{i|\mathbf{k}|r \cos \phi} d\phi dr \\ &= 4\pi \int_0^\infty R(r) r \frac{\sin |\mathbf{k}|r}{|\mathbf{k}|} dr \end{aligned}$$

Since the correlation function $R(\cdot)$ is defined only on $[0, +\infty)$, we extend it symmetrically to the whole real line by $R(-\cdot) = R(\cdot)$ for any value in $(-\infty, 0)$. Then, the above integral becomes

$$\widehat{R}(|\mathbf{k}|) = \frac{2\pi}{|\mathbf{k}|} \int_{-\infty}^\infty R(r) r \sin |\mathbf{k}|r dr = -\frac{2\pi}{|\mathbf{k}|^3} \int_{-\infty}^\infty \dot{R}(r) (\sin |\mathbf{k}|r - |\mathbf{k}|r \cos |\mathbf{k}|r) dr,$$

where the last equality holds due to integration by parts and the exponential decay of $R(\cdot)$. Then, with Riemann Lebesgue lemma, the following limit is

$$\lim_{K \rightarrow \infty} \int_0^K |\mathbf{k}|^3 \widehat{R}(|\mathbf{k}|) d|\mathbf{k}| = -2\pi \int_{-\infty}^\infty \frac{\dot{R}(r)}{r} dr.$$

The symmetric extension of $R(\cdot)$ implies that $\dot{R}(r) = -\dot{R}(-r)$ when $r < 0$. Therefore,

$$\int_0^\infty r^3 \widehat{R}(r) dr = \lim_{K \rightarrow \infty} \int_0^K |\mathbf{k}|^3 \widehat{R}(|\mathbf{k}|) d|\mathbf{k}| = -4\pi \int_0^\infty \frac{\dot{R}(r)}{r} dr,$$

which implies that the integral on the right hand side is not positive. ■

Note that the conditions on the autocorrelation function R in Lemma B.1 are not restrictive. Many autocorrelation functions satisfy those conditions, as for example, $R(|\mathbf{x}|) = e^{-\frac{1}{2}|\mathbf{x}|^2}$ for a Gaussian random field and the power law function $R(|\mathbf{x}|) = (1 + |\mathbf{x}|)e^{-|\mathbf{x}|}$. With this lemma, we can now prove Proposition 5.1.

Proof of Proposition 5.1. To simplify the notation, we first define the path integral of the random process function $\mu(\cdot)$ as

$$\nu(\mathbf{x}, \mathbf{y}) = \int_0^1 \mu\left(\frac{\mathbf{x}}{l} + s\frac{\mathbf{y} - \mathbf{x}}{l}\right) ds,$$

and

$$\vartheta(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| \nu(\mathbf{x}, \mathbf{y}).$$

We need to compute moment estimation

$$\mathbb{E}(\widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)}) = \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} E,$$

where

$$E = \mathbb{E}\left(e^{-i\kappa\sigma(|\mathbf{x} - \mathbf{y}_1| \nu(\mathbf{x}, \mathbf{y}_1) - |\mathbf{x} - \mathbf{y}_2| \nu(\mathbf{x}, \mathbf{y}_2))}\right) = e^{-\frac{1}{2}\kappa^2\sigma^2\mathbb{E}(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2))^2}. \quad (54)$$

Because $|\mathbf{y}_1 - \mathbf{y}_2| \ll L$ we can write

$$|\mathbf{x} - \mathbf{y}_2| = |\mathbf{x} - \mathbf{y}_1| + \frac{(\mathbf{x} - \mathbf{y}_1)^T}{|\mathbf{x} - \mathbf{y}_1|}(\mathbf{y}_2 - \mathbf{y}_1) + o(|\mathbf{y}_2 - \mathbf{y}_1|),$$

and then

$$\nu(\mathbf{x}, \mathbf{y}_2) = \nu(\mathbf{x}, \mathbf{y}_1) + \nabla_{\mathbf{y}_1}^T \nu(\mathbf{x}, \mathbf{y}_1)(\mathbf{y}_2 - \mathbf{y}_1) + o(|\mathbf{y}_2 - \mathbf{y}_1|).$$

Dropping terms of order 2 or higher, we estimate the expectation in the exponent as

$$\begin{aligned} & \mathbb{E}\left(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2)\right)^2 \\ & \approx \mathbb{E}\left(\nu(\mathbf{x}, \mathbf{y}_1) \frac{(\mathbf{x} - \mathbf{y}_1)^T}{|\mathbf{x} - \mathbf{y}_1|}(\mathbf{y}_2 - \mathbf{y}_1) + |\mathbf{x} - \mathbf{y}_1| \nabla_{\mathbf{y}_1}^T \nu(\mathbf{x}, \mathbf{y}_1)(\mathbf{y}_2 - \mathbf{y}_1)\right)^2. \end{aligned}$$

For simplicity, we consider \mathbf{y}_1 and \mathbf{y}_2 on the same plane, at a distance L from point $\mathbf{x} = (0, 0, 0)$, so $\mathbf{y}_1 = (0, 0, L)$ and $\mathbf{y}_2 = (\xi, \eta, L)$. Then, $(\mathbf{y}_2 - \mathbf{y}_1)^T(\mathbf{x} - \mathbf{y}_1) = 0$ and the variance is reduced to the estimation of

$$\begin{aligned} & \mathbb{E}\left(\nabla_{\mathbf{y}}^T \nu(\mathbf{x}, \mathbf{y})(\mathbf{y}' - \mathbf{y})\right)^2 \\ & = \xi^2 \mathbb{E}\left(\frac{\partial}{\partial y_1} \nu(\mathbf{x}, \mathbf{y})\right)^2 + \eta^2 \mathbb{E}\left(\frac{\partial}{\partial y_2} \nu(\mathbf{x}, \mathbf{y})\right)^2 + 2\xi\eta \mathbb{E}\left(\frac{\partial}{\partial y_1} \nu(\mathbf{x}, \mathbf{y}) \frac{\partial}{\partial y_2} \nu(\mathbf{x}, \mathbf{y})\right) \\ & = \left(\xi^2 \frac{\partial}{\partial y_1} \frac{\partial}{\partial y'_1} \mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}')) + \eta^2 \frac{\partial}{\partial y_2} \frac{\partial}{\partial y'_2} \mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}')) \right. \\ & \quad \left. + 2\xi\eta \frac{\partial}{\partial y_1} \frac{\partial}{\partial y'_2} \mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}'))\right) \Big|_{\mathbf{y}' = \mathbf{y}}, \end{aligned} \quad (55)$$

where $\frac{\partial}{\partial y_i}$ means the derivative with respect to the i^{th} component of vector \mathbf{y} .

Denote $\tilde{\mathbf{y}}' = [\xi, \eta]^T$ and $\tilde{\mathbf{y}} = [0, 0]^T$, such that $\mathbf{y} = [\tilde{\mathbf{y}}^T, L]^T$ and $\mathbf{y}' = [\tilde{\mathbf{y}}'^T, L]^T$, and let $\varsigma = \sqrt{\frac{1}{l^2}|s\tilde{\mathbf{y}} - s'\tilde{\mathbf{y}}'|^2 + \frac{L^2}{l^2}(s - s')^2}$. We compute the derivatives of the expectations

$$\mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}')) = \int_0^1 \int_0^1 R(\varsigma) ds' ds$$

in (55) as follows:

$$\begin{aligned}
\frac{\partial}{\partial y_1} \mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}')) &= \int_0^1 \int_0^1 \dot{R}(\varsigma) \frac{s}{\varsigma l^2} (sy_1 - s'y'_1) ds' ds, \\
\frac{\partial^2}{\partial y'_1 \partial y_1} \mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}')) &= \int_0^1 \int_0^1 \left(\dot{R}(\varsigma) \frac{ss'(sy_1 - s'y'_1)^2}{\varsigma^3 l^4} - \dot{R}(\varsigma) \frac{ss'}{\varsigma l^2} - \ddot{R}(\varsigma) \frac{ss'}{\varsigma^2 l^4} (sy_1 - s'y'_1)^2 \right) ds' ds, \\
\frac{\partial^2}{\partial y_1 \partial y'_2} \mathbb{E}(\nu(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, \mathbf{y}')) &= \int_0^1 \int_0^1 \frac{ss'}{\varsigma^2 l^4} \left(\frac{\dot{R}(\varsigma)}{\varsigma} - \ddot{R}(\varsigma) \right) (sy_1 - s'y'_1)(sy_2 - s'y'_2) ds' ds.
\end{aligned}$$

Taking $\mathbf{y} = \mathbf{y}' = (0, 0, L)$ in the above second order derivatives, we obtain that

$$\mathbb{E} \left(\frac{\partial}{\partial y_1} \nu(\mathbf{x}, \mathbf{y}) \frac{\partial}{\partial y_2} \nu(\mathbf{x}, \mathbf{y}) \right) = 0,$$

so that

$$\mathbb{E} \left(\nabla_{\mathbf{y}}^T \nu(\mathbf{x}, \mathbf{y}) (\mathbf{y}' - \mathbf{y}) \right)^2 = (\xi^2 + \eta^2) \mathbb{E} \left(\frac{\partial}{\partial y_1} \nu(\mathbf{x}, \mathbf{y}) \right)^2.$$

Next, we compute

$$\begin{aligned}
\mathbb{E} \left(\frac{\partial}{\partial y_1} \nu(\mathbf{x}, \mathbf{y}) \right)^2 &= - \int_0^1 \int_0^1 \dot{R} \left(\frac{L|s - s'|}{l} \right) \frac{ss'}{lL|s - s'|} ds' ds \\
&= \frac{2l^2}{3L^4} \int_0^{L/l} sR(s) ds - \frac{2}{3lL} \int_0^{L/l} \frac{\dot{R}(t)}{t} dt + \frac{2}{3L^2} R \left(\frac{L}{l} \right) - \frac{R(0)}{L^2} \\
&\approx \frac{1}{L^2} \left(-1 - \frac{2L}{3l} \int_0^\infty \frac{\dot{R}(t)}{t} dt \right),
\end{aligned}$$

where the last approximation is based on the condition $l \ll L$ and $R(\cdot)$ has exponential decay with normalization $R(0) = 1$. Now, let $a_e^2 = \sigma^2 L^4 \mathbb{E} \left(\frac{\partial}{\partial y_1} \nu(\mathbf{x}, \mathbf{y}) \right)^2$. We then have

$$\mathbb{E}(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2))^2 \approx |\mathbf{x} - \mathbf{y}_1|^2 \mathbb{E}(\nabla_{\mathbf{y}}^T \nu(\mathbf{x}, \mathbf{y}_1) (\mathbf{y}_2 - \mathbf{y}_1))^2 \approx \frac{|\mathbf{y}_1 - \mathbf{y}_2|^2}{\sigma^2 L^2} a_e^2, \quad (56)$$

and using (54)

$$\mathbb{E}(\widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)}) = \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} E \approx \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2},$$

i.e. the moment estimate (36). ■

Proof of Corollary 5.2. The result is a direct application of Proposition 5.1. Let $\mathcal{C} = \frac{1}{(4\pi|\mathbf{x}-\mathbf{y}_1|)^2(4\pi|\mathbf{x}-\mathbf{y}_2|)^2}$. Using (36), we have

$$\begin{aligned}
& \mathbb{E} \left| \widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)} - \mathbb{E} \left(\widehat{G}(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}(\mathbf{x}, \mathbf{y}_2)} \right) \right|^2 \\
& \approx \mathbb{E} \left| \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} \left(e^{i\kappa\sigma(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2))} - e^{-\frac{1}{2}\kappa^2 \frac{a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} \right) \right|^2 \\
& = \mathcal{C} \mathbb{E} \left(1 + e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} - e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} (e^{i\kappa\sigma(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2))} + e^{-i\kappa\sigma(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2))}) \right) \\
& = \mathcal{C} \left(1 + e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} - 2e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} e^{-\frac{1}{2}\kappa^2 \sigma^2 \mathbb{E}(\vartheta(\mathbf{x}, \mathbf{y}_1) - \vartheta(\mathbf{x}, \mathbf{y}_2))^2} \right) \\
& \approx \mathcal{C} (1 - e^{-\kappa^2 \frac{a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2}),
\end{aligned}$$

where we use the estimate (56) in the last step. ■

Proof of Proposition 5.3. For simplicity, we use the same configuration as in the proof of Proposition 5.1, where $\mathbf{y}_1 = [0, 0, L]^T$ and $\mathbf{y}_2 = [\xi, \eta, L]^T$. We first look at the denominator of the ratio in (39)

$$\|\widehat{\mathbf{g}}(\mathbf{y}_1)\|^2 = \sum_{j=1}^N \left| \widehat{G}(\mathbf{x}_j, \mathbf{y}_1) \right|^2 = \sum_{j=1}^N \left| \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1) \right|^2 = \|\widehat{\mathbf{g}}_0(\mathbf{y}_1)\|^2. \quad (57)$$

Using the same approach as in the proof in [16], under continuous limit, we have

$$\|\widehat{\mathbf{g}}(\mathbf{y}_1)\|^2 \approx \frac{1}{(4\pi h)^2} \int_{\Omega(\mathbf{x})} \frac{d\mathbf{x}}{|\mathbf{x} - \mathbf{y}|^2} = \frac{1}{16\pi h^2} \log \left(1 + \frac{a^2}{4L^2} \right).$$

On the other hand, the numerator can be computed as follows

$$\begin{aligned}
& \mathbb{E} |\widehat{\mathbf{g}}^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E} (\widehat{\mathbf{g}}^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2))|^2 \\
& = \mathbb{E} \left| \sum_{j=1}^N \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2)} \left(e^{i\kappa\sigma(\vartheta(\mathbf{x}_j, \mathbf{y}_1) - \vartheta(\mathbf{x}_j, \mathbf{y}_2))} - e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} \right) \right|^2 \\
& = \sum_{j, j'=1}^N \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2)} \widehat{G}_0(\mathbf{x}_{j'}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}_{j'}, \mathbf{y}_2)} \\
& \quad \times \mathbb{E} \left(e^{i\kappa\sigma(\vartheta(\mathbf{x}_j, \mathbf{y}_1) - \vartheta(\mathbf{x}_j, \mathbf{y}_2) - \vartheta(\mathbf{x}_{j'}, \mathbf{y}_1) + \vartheta(\mathbf{x}_{j'}, \mathbf{y}_2))} \right. \\
& \quad \left. - e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} e^{-i\kappa\sigma(\vartheta(\mathbf{x}_{j'}, \mathbf{y}_1) - \vartheta(\mathbf{x}_{j'}, \mathbf{y}_2))} \right. \\
& \quad \left. - e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} e^{-i\kappa\sigma(\vartheta(\mathbf{x}_j, \mathbf{y}_1) - \vartheta(\mathbf{x}_j, \mathbf{y}_2))} + e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} \right) \\
& = \sum_{j, j'=1}^N \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2)} \widehat{G}_0(\mathbf{x}_{j'}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}_{j'}, \mathbf{y}_2)} \\
& \quad \times e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2} \left(e^{\kappa^2 \sigma^2 \mathbb{E}(\vartheta(\mathbf{x}_j, \mathbf{y}_1) - \vartheta(\mathbf{x}_j, \mathbf{y}_2))(\vartheta(\mathbf{x}_{j'}, \mathbf{y}_1) - \vartheta(\mathbf{x}_{j'}, \mathbf{y}_2))} - 1 \right).
\end{aligned}$$

From the above expression, we can see the variance of $\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2)$ is close to the denominator up to the factor

$$e^{-\frac{\kappa^2 a_e^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2} \left(e^{\kappa^2 \sigma^2 \mathbb{E}(\vartheta(\mathbf{x}_j, \mathbf{y}_1) - \vartheta(\mathbf{x}_j, \mathbf{y}_2))(\vartheta(\mathbf{x}_{j'}, \mathbf{y}_1) - \vartheta(\mathbf{x}_{j'}, \mathbf{y}_2))} - 1 \right). \quad (58)$$

The expectation

$$\mathbb{E}(\vartheta(\mathbf{x}_j, \mathbf{y}_1) - \vartheta(\mathbf{x}_j, \mathbf{y}_2))(\vartheta(\mathbf{x}_{j'}, \mathbf{y}_1) - \vartheta(\mathbf{x}_{j'}, \mathbf{y}_2))$$

is nonzero only when there is strong correlation between the path starting from \mathbf{x}_j and $\mathbf{x}_{j'}$. This is controlled by the correlation length l of the random medium. When those paths are within l , the value of (58) is reduced to $1 - e^{-\frac{\kappa^2 a_e^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2}$, and otherwise (58) is equal to zero. Thus in the continuous limit, the value of numerator can be approximated by

$$\begin{aligned} \mathbb{E}|\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2))|^2 &\approx \frac{1}{h^4} \left(1 - e^{-\frac{\kappa^2 a_e^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2} \right) \\ &\times \int_{\Omega(\mathbf{x})} \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} d\mathbf{x} \int_{B(\mathbf{0}, l)} \widehat{G}_0(\mathbf{x}', \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}', \mathbf{y}_2)} d\mathbf{x}', \end{aligned} \quad (59)$$

where $B(\mathbf{0}, l)$ is a ball centered at $\mathbf{0}$ with radius l within which factor (58) is not equal to zero.

Therefore the ratio on the left handside of (39) is bounded by

$$\begin{aligned} &\left| \frac{\mathbb{E}|\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2))|^2}{\|\widehat{\mathbf{g}}(\mathbf{y}_1)\|^2 \|\widehat{\mathbf{g}}(\mathbf{y}_2)\|^2} \right| \\ &\approx \frac{256\pi^2 \left(1 - e^{-\frac{\kappa^2 a_e^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2} \right)}{\log^2 \left(1 + \frac{a^2}{4L^2} \right)} \left| \int_{\Omega(\mathbf{x})} \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} d\mathbf{x} \right| \\ &\quad \times \left| \int_{B(\mathbf{0}, l)} \widehat{G}_0(\mathbf{x}', \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}', \mathbf{y}_2)} d\mathbf{x}' \right| \\ &\leq \frac{\left(1 - e^{-\frac{\kappa^2 a_e^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2} \right)}{\log^2 \left(1 + \frac{a^2}{4L^2} \right)} \int_{B(\mathbf{0}, l)} \frac{d\mathbf{x}}{|\mathbf{x} - \mathbf{y}|^2} \int_{\Omega(\mathbf{x})} \frac{d\mathbf{x}}{|\mathbf{x} - \mathbf{y}|^2} \\ &\approx \frac{l^2 \left(1 - e^{-\frac{\kappa^2 a_e^2}{L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2} \right)}{L^2 \log \left(1 + \frac{a^2}{4L^2} \right)}. \end{aligned}$$

When size of array a increases, we have $\frac{\mathbb{E}|\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2))|^2}{\|\widehat{\mathbf{g}}(\mathbf{y}_1)\|^2 \|\widehat{\mathbf{g}}(\mathbf{y}_2)\|^2}$ goes to 0 logarithmically. ■

Remark B.2. The result in Proposition 5.3 holds for any regime no matter $a \ll L$ or $L \lesssim a$. When $a \ll L$, i.e. in the paraxial regime, we can use parabolic approximation to compute the approximate value for the ratio on the lefthand side of (39). Let $\mathbf{x} = [x_1, x_2, 0]^T$. Then we have

$$|\mathbf{x} - \mathbf{y}_1| = \sqrt{x_1^2 + x_2^2 + L^2} \approx L + \frac{x_1^2 + x_2^2}{2L}$$

and

$$|\mathbf{x} - \mathbf{y}_2| = \sqrt{L^2 + (x_1 - \xi)^2 + (x_2 - \eta)^2} = L + \frac{(x_1 - \xi)^2 + (x_2 - \eta)^2}{2L}.$$

Since $|\mathbf{y}_1 - \mathbf{y}_2| \leq a \ll L$, we approximate in denominator

$$|\mathbf{x} - \mathbf{y}_1| \approx |\mathbf{x} - \mathbf{y}_2| \approx L,$$

and for the two integrals in (59),

$$\begin{aligned} \int_{\Omega(\mathbf{x})} \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} d\mathbf{x} &\approx \frac{e^{-i\kappa \frac{\xi^2 + \eta^2}{2L}}}{L^2} \int_{\Omega(\mathbf{x})} e^{i\kappa \frac{x_1 \xi + x_2 \eta}{L}} d\mathbf{x} \\ &= \frac{e^{-i\kappa \frac{\xi^2 + \eta^2}{2L}}}{L^2} \int_{-a/2}^{a/2} e^{i\kappa \frac{x_1 \xi}{L}} dx_1 \int_{-a/2}^{a/2} e^{i\kappa \frac{x_2 \eta}{L}} dx_2 \\ &= \frac{a^2 e^{-i\kappa \frac{\xi^2 + \eta^2}{2L}}}{\pi L^2} \text{sinc}\left(\frac{a\xi}{\lambda L}\right) \text{sinc}\left(\frac{a\eta}{\lambda L}\right), \end{aligned}$$

and similarly,

$$\int_{B(\mathbf{0}, l)} \widehat{G}_0(\mathbf{x}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_2)} d\mathbf{x} \approx \frac{a^2 e^{-i\kappa \frac{\xi^2 + \eta^2}{2L}}}{\pi L^2} \text{sinc}\left(\frac{l\xi}{\lambda L}\right) \text{sinc}\left(\frac{l\eta}{\lambda L}\right).$$

Note that when $a \ll L$, we can approximate linearly $\log\left(1 + \frac{a^2}{4L^2}\right) \approx \frac{a^2}{4L^2}$. Therefore, the lefthand side of (39) is approximately equal to

$$\begin{aligned} &\frac{\mathbb{E} |\widehat{\mathbf{g}}^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2))|^2}{\|\widehat{\mathbf{g}}(\mathbf{y}_1)\|^2 \|\widehat{\mathbf{g}}(\mathbf{y}_2)\|^2} \\ &\approx \frac{a^2 l^2 \left(1 - e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2}\right) \text{sinc}\left(\frac{\xi a}{\lambda L}\right) \text{sinc}\left(\frac{\eta a}{\lambda L}\right) \text{sinc}\left(\frac{\xi l}{\lambda L}\right) \text{sinc}\left(\frac{\eta l}{\lambda L}\right)}{\frac{L^4 \pi^2}{256 \pi^2} \frac{a^2}{4L^2} \frac{a^2}{4L^2}} \\ &\approx 256 \pi^2 \left(1 - e^{-\frac{\kappa^2 a_e^2}{L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2}\right) \left(\frac{l}{a}\right)^2 \text{sinc}\left(\frac{\xi a}{\lambda L}\right) \text{sinc}\left(\frac{\eta a}{\lambda L}\right) \text{sinc}\left(\frac{\xi l}{\lambda L}\right) \text{sinc}\left(\frac{\eta l}{\lambda L}\right). \end{aligned} \tag{60}$$

This matches the result of logarithmic decay bound in general.

Proof of Proposition 5.4. We first look at (42) for back-propagation in the true random media scenario. It has been shown in Proposition 5.3 that when size of array is large, the single realization will approach the average value. Therefore, it is enough to show the average value goes to zero as points \mathbf{y}_1 and \mathbf{y}_2 are far apart and the result is true due to the Chebyshev inequality under probability measure \mathbb{P} induced from random field $\mu(\cdot)$. Using the moment formula, we have

$$\mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2)) = \sum_{j=1}^N \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2) e^{-\frac{\kappa^2 a_e^2}{2L^2} |\mathbf{y}_1 - \mathbf{y}_2|^2}.$$

Since the multiplier factor $e^{-\frac{\kappa^2 a_e^2}{2L^2}|\mathbf{y}_1 - \mathbf{y}_2|^2} < 1$ and goes to zero as $\frac{|\mathbf{y}_1 - \mathbf{y}_2|}{\lambda}$ increases, we thus have

$$\mathbb{E}(\widehat{\mathbf{g}}^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2)) \rightarrow 0, \quad \frac{|\mathbf{y}_1 - \mathbf{y}_2|}{\lambda} \rightarrow \infty$$

especially when size of array a becomes large. When a_e is large in random media, average value will decay to zero faster than that in homogeneous media. This is why in random media better resolution can be achieved.

Next, we look at (43). First, we compute the expectation of $\nu^2(\mathbf{x}, \mathbf{y})$

$$\mathbb{E}\nu^2(\mathbf{x}, \mathbf{y}) = 2 \int_0^1 \int_0^s R\left(\frac{|\mathbf{y} - \mathbf{x}|}{l}(s - s')\right) ds' ds = \frac{2l}{|\mathbf{y} - \mathbf{x}|} \int_0^1 \int_0^{\frac{|\mathbf{y} - \mathbf{x}|}{l}s} R(s') ds' ds$$

Under $l \ll L$ and symmetric extension of $R(\cdot)$ to negative real line, we have

$$\mathbb{E}\nu^2(\mathbf{x}, \mathbf{y}) \approx \frac{2l}{|\mathbf{y} - \mathbf{x}|} \int_0^\infty R(s) ds = \frac{l\widehat{R}(0)}{|\mathbf{y} - \mathbf{x}|}.$$

Then it is easy to calculate the expectation of mixed inner product as

$$\begin{aligned} \mathbb{E}(\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2)) &= \sum_{j=1}^N \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2) \mathbb{E}e^{i\kappa\sigma|\mathbf{x}_j - \mathbf{y}_2|\nu(\mathbf{x}_j, \mathbf{y}_2)} \\ &= \sum_{j=1}^N \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2) e^{-\frac{1}{2}\kappa^2\sigma_0^2|\mathbf{x}_j - \mathbf{y}_2|^2\mathbb{E}\nu^2(\mathbf{x}_j, \mathbf{y}_2)} \\ &\approx \sum_{j=1}^N \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2) e^{-\frac{1}{2}\kappa^2\sigma_0^2 l \widehat{R}(0)|\mathbf{x}_j - \mathbf{y}_2|}. \end{aligned}$$

Because $\widehat{R}(0) \geq 0$, it implies $\mathbb{E}(\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2))$ is approximately equal to $\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}_0(\mathbf{y}_2)$ multiplied by a factor less than 1 and therefore

$$\mathbb{E}(\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2)) \rightarrow 0, \quad \frac{|\mathbf{y}_1 - \mathbf{y}_2|}{\lambda} \rightarrow \infty.$$

To show the statistical stability of the mixed inner product, we first compute the bound of the numerator similar to that in the proof of Proposition 5.4.

$$\begin{aligned} &\mathbb{E}\left|\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}\widehat{\mathbf{g}}_0^*(\mathbf{y}_1)\widehat{\mathbf{g}}(\mathbf{y}_2)\right|^2 \\ &\approx \mathbb{E}\left|\sum_{j=1}^N \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2) \left(e^{i\kappa\sigma\vartheta(\mathbf{x}_j, \mathbf{y}_2)} - e^{-\frac{1}{2}\kappa^2\sigma^2 l \widehat{R}(0)|\mathbf{x}_j - \mathbf{y}_2|}\right)\right|^2 \\ &= \sum_{j, j'=1}^N \overline{\widehat{G}_0(\mathbf{x}_j, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}_j, \mathbf{y}_2) \widehat{G}_0(\mathbf{x}_{j'}, \mathbf{y}_1) \overline{\widehat{G}_0(\mathbf{x}_{j'}, \mathbf{y}_2)} \\ &\quad \times e^{-\frac{1}{2}\kappa^2\sigma^2 l \widehat{R}(0)(|\mathbf{x}_j - \mathbf{y}_2| + |\mathbf{x}_{j'} - \mathbf{y}_2|)} \left(e^{\kappa^2\sigma^2 \mathbb{E}\vartheta(\mathbf{x}, \mathbf{y}_2)\vartheta(\mathbf{x}_{j'}, \mathbf{y}_2)} - 1\right). \end{aligned}$$

The correlation term $\vartheta(\mathbf{x}_j, \mathbf{y}_2)\vartheta(\mathbf{x}_{j'}, \mathbf{y}_2)$ is nonzero only when the paths connecting $\mathbf{x}_j, \mathbf{x}_{j'}$ with \mathbf{y}_2 are within the correlation length l . Also according to Cauchy-Schwartz inequality,

$$\mathbb{E}(\vartheta(\mathbf{x}_j, \mathbf{y}_2)\vartheta(\mathbf{x}_{j'}, \mathbf{y}_2)) \leq \frac{1}{2} (\mathbb{E}\vartheta^2(\mathbf{x}_j, \mathbf{y}_2) + \mathbb{E}\vartheta^2(\mathbf{x}_{j'}, \mathbf{y}_2)) .$$

Therefore the numerator is approximately bounded by continuous limit

$$\begin{aligned} & \mathbb{E} \left| \widehat{\mathbf{g}}_0^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E} \widehat{\mathbf{g}}_0^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2) \right|^2 \\ & \lesssim \frac{1}{h^4} \left(1 - e^{-\kappa^2 \sigma^2 L l \widehat{R}(0)} \right) \left| \int_{\Omega(\mathbf{x})} \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}, \mathbf{y}_2) d\mathbf{x} \right| \left| \int_{B(\mathbf{0}, l)} \overline{\widehat{G}_0(\mathbf{x}, \mathbf{y}_1)} \widehat{G}_0(\mathbf{x}, \mathbf{y}_2) d\mathbf{x} \right| \\ & \lesssim \frac{1}{256\pi^2 h^4} \left(1 - e^{-\kappa^2 \sigma^2 L l \widehat{R}(0)} \right) \log \left(1 + \frac{a^2}{4L^2} \right) \log \left(1 + \frac{l^2}{4L^2} \right). \end{aligned}$$

The denominator is norm of the Green function vector which has been calculated in the proof of Proposition 5.3. Thus the ratio in (44) is bounded by

$$\frac{\mathbb{E} |\widehat{\mathbf{g}}_0^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2) - \mathbb{E}(\widehat{\mathbf{g}}_0^*(\mathbf{y}_1) \widehat{\mathbf{g}}(\mathbf{y}_2))|^2}{\|\widehat{\mathbf{g}}_0(\mathbf{y}_1)\|^2 \|\widehat{\mathbf{g}}(\mathbf{y}_2)\|^2} \lesssim \left(1 - e^{-\kappa^2 \sigma^2 L l \widehat{R}(0)} \right) \frac{\log \left(1 + \frac{l^2}{4L^2} \right)}{\log \left(1 + \frac{a^2}{4L^2} \right)}.$$

When $a \rightarrow \infty$, the right handside goes to zero. The decay rate is again controlled by the logarithmic factor of a and when $a \ll L$, the decay rate is quadratic of a which is the same as that in Proposition 5.3. Due to the statistical stability, the mixed inner product satisfies (43). ■

References

- [1] Baggeroer A, Kuperman W and Mikhalevsky P 1993 An overview of matched field methods in ocean acoustics *IEEE Journal of Oceanic Engineering* **18** 401–24
- [2] Bal G, Papanicolaou G and Ryzhik L 2002 Self-averaging in time reversal for the parabolic wave equation *Stochastics and Dynamics* **2** 507–531
- [3] Baraniuk R and Steeghs P 2007 Compressive radar imaging Radar Conference, 2007 IEEE 128–33
- [4] Biondi B 2006 *3D seismic imaging* (Society of Exploration Geophysicists **14**)
- [5] Blomgren P, Papanicolaou G and Zhao H 2002 Super-Resolution in Time-Reversal Acoustics *J. Acoust. Soc. Am.* **111** 230–48
- [6] Bobin J, Starck J and Ottensamer R 2008 Compressed Sensing in Astronomy *IEEE Journal of Selected Topics in Signal Processing* **2** 718–26
- [7] Borcea L, Papanicolaou G and Tsogka C 2005 Interferometric array imaging in clutter *Inverse Problems* **21** 1419–60

- [8] Borcea L, Papanicolaou G and Tsogka C 2006 Adaptive interferometric imaging in clutter and optimal illumination *Inverse Problems* **22** 1405–36
- [9] Borcea L, Garnier J, Papanicolaou G and Tsogka C 2011 Enhance statistical stability in coherent interferometric imaging *Inverse Problems* **27** 085004
- [10] Borcea L and Kocyigit I 2015 Resolution analysis of imaging with ℓ_1 optimization *SIAM J. Imaging Sci.* in press
- [11] Candès E, Romberg J and Tao T 2006 Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52(2)** 489–509
- [12] Candès EJ and Tao T 2006 Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inf. Theory* **52(12)** 5406–25
- [13] Candès EJ, Eldar YC, Strohmer T, and Voroninski V 2013 Phase retrieval via matrix completion *SIAM J. Imaging Sci.* **6** 199–225
- [14] Candès EJ, Strohmer T, and Voroninski V 2013 PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming *Communications on Pure and Applied Mathematics* **66** 1241–1274
- [15] Chai A, Moscoso M and Papanicolaou G 2011 Array imaging using intensity-only measurements *Inverse Problems* **27(1)** 015005
- [16] Chai A, Moscoso M and Papanicolaou G 2013 Robust imaging of localized scatterers using the singular value decomposition and ℓ_1 minimization *Inverse Problems* **29(2)** 025016
- [17] Chai A, Moscoso M and Papanicolaou G 2014 Imaging strong localized scatterers with sparsity promoting optimization *SIAM J. Imaging Sci.* **7(2)** 1358–87
- [18] Donoho D and Stark P 1989 Uncertainty principles and signal recovery *SIAM J. Appl. Math.* **49(3)** 906–31
- [19] Donoho D and Logan B 1992 Signal recovery and the large sieve *SIAM J. Appl. Math.* **52(2)** 577–91
- [20] Donoho D 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52(4)** 1289–1306
- [21] Fannjiang A, Strohmer T and Yan P 2010 Compressed remote sensing of sparse objects *SIAM J. Imaging Sci.* **3** 595–618
- [22] Fannjiang A 2012 Uniqueness of phase retrieval with random illumination *Inverse Problems* **28** 075008 (2012)
- [23] Fannjiang A and Liao W 2012 Phase retrieval with random phase illumination *J. Opt. Soc. A* **29** 1847–1859

- [24] Fannjiang A and Tseng HC 2013 Compressive radar with off-grid targets: a perturbation approach *Inverse Problems* **29** 054008
- [25] Fink M 1993 Time reversal mirrors *J. Phys. D: Appl. Phys.* **26** 1330–50
- [26] Foldy L 1945 The multiple scattering of waves *Pyhs. Rev.* **67** 107–19
- [27] Lax M 1951 Multiple scattering of waves *Rev. Modern. Phys.* **23** 287–310
- [28] Lax M 1952 Multiple scattering of waves II, The effective field in dense systems *Phys. Rev.* **85** 261–9
- [29] Lustig M, Donoho D and Pauly J 2007 Sparse MRI: The application of compressed sensing for rapid MR imaging *Magn Reson Med* **58(6)** 1182–95
- [30] Moscoso M, Novikov A Papanicolaou G and Ryzhik L 2012 A differential equations approach to ℓ_1 -minimization with applications to array imaging *Inverse Problems* **28(10)** 105001
- [31] Moscoso M, Novikov A and Papanicolaou G 2015 Coherent imaging without phases *arXiv preprint* arXiv:1510.04158
- [32] Novikov A, Moscoso M and Papanicolaou G 2015 Illumination strategies for intensity-only imaging *SIAM J. Imaging Sci.* **8(3)** 1547–73
- [33] Papanicolaou G, Ryzhik L and Solna K 2004 Statistical stability in time reversal *SIAM J. Appl. Math.* **64** 1133–55
- [34] Prada C, Manneville S, Spolianski D and Fink M 1996 Decomposition of the time reversal operator: detection and selective focusing on two scattereres *J. Acoust. Soc. Am.* **99** 2067–76
- [35] Rytov S M, Kravtsov Y A, Tatarskii V I 1989 *Principles of statistical radiophysics. 4. Wave Propagation through random media* (Springer Verlag, Berlin)
- [36] Schmidt R 1986 Multiple emitter location and signal parameter estimation *IEEE Transactions on Antennas and Propagation* **34(3)** 276–80
- [37] Studer V, Bobin J, Chahid M, Moussavi H, Candès E and Dahan M 2012 Compressive fluorescence microscopy for biological and hyperspectral imaging *Proceedings of the National Academy of Sciences USA* **109(26)** E1679–87
- [38] Tatarski V I 1961 *Wave propagation in a turbulent medium* (Dover, New York)
- [39] Taylor H, Banks S and McCoy J 1979 Deconvolution with the ℓ_1 norm *Geophysics* **44(1)** 39–52
- [40] Wu Y, Ye P, Mirza I, Arce G and Prather D 2009 Experimental demonstration of anoptical-sectioning compressive sensing microscope (CSM) *Opt. Express* **18** 24565–78